

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2004, 12:48:09 ; Search time 55 Seconds
(without alignments)
82.196 Million cell updates/sec

Title: US-09-933-780C-16
Perfect score: 92
Sequence: 1 SRHHCRSKAKRSRHH 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	92	100.0	16	5 AAU78950	Aau78950 Human PER
2	92	100.0	16	5 AAU78934	Aau78934 Human per
3	92	100.0	16	7 ADB39032	Adb39032 Human tra
4	92	100.0	16	7 ADB39085	Adel3829 Human per
5	92	100.0	25	7 ADB39085	Adb39085 Human hPE
6	92	100.0	25	7 ADB39082	Adb39082 Human hPE
7	92	100.0	40	5 AAU78964	Aau78964 Human per
8	92	100.0	70	5 AAU78965	Aau78965 Human per
9	92	100.0	1290	2 AAU78968	Aay06809 Human per
10	92	100.0	1290	3 AAU78968	Aay32216 Human per
11	92	100.0	1290	5 ABB09289	Abb09289 Human per
12	92	100.0	1291	2 AAU78961	Aay01687 Protein e
13	89	96.7	16	5 AAU78961	Aau78961 Human PER
14	89	96.7	16	5 AAU78957	Aau78957 Human PER
15	88	95.7	16	5 AAU78926	Aau78926 Human per
16	88	95.7	16	5 AAU78966	Aau78966 Mouse per
17	88	95.7	1291	2 AAU78961	Aay01689 Protein e
18	88	95.7	1291	2 AAU78961	Aay06810 Mouse per
19	86	93.5	16	5 AAU78959	Aau78959 Human PER
20	86	93.5	16	5 AAU78952	Aau78952 Human PER
21	86	93.5	16	5 AAU78958	Aau78958 Human PER
22	86	93.5	16	5 AAU78956	Aau78956 Human PER
23	86	93.5	16	5 AAU78951	Aau78951 Human PER
24	86	93.5	16	5 AAU78960	Aau78960 Human PER
25	86	93.5	16	5 AAU78962	Aau78962 Human PER

26	83	90.2	16	5 AAU78955	Aau78955 Human PER
27	82	89.1	16	5 AAU78954	Aau78954 Human PER
28	82	89.1	16	5 AAU78953	Aau78953 Human PER
29	81	88.0	18	6 ABU08204	Abu08204 Human per
30	79	85.9	26	5 AAU78927	Aau78927 Flag tagg
31	72	78.3	13	5 AAU78963	Aau78963 Human tra
32	60	65.2	14	7 ADB39033	Adb39033 Human hPE
33	60	65.2	22	7 ADB39086	Adb39086 Human hPE
34	60	65.2	22	7 ADB39079	Adb39079 Human hPE
35	60	65.2	22	7 ADB39083	Adb39083 Human hPE
36	60	65.2	26	7 ADB39081	Adb39081 Human hPE
37	60	65.2	26	7 ADB39080	Adb39080 Human hPE
38	60	65.2	27	7 ADB39087	Adb39087 Human hPE
39	60	65.2	28	7 ADB39084	Adb39084 Human hPE
40	60	65.2	32	7 ADB39090	Adb39090 Human hPE
41	60	65.2	37	7 ADB39091	Adb39091 Human hPE
42	51.5	56.0	375	4 ABB64168	Abb64168 Drosophil
43	51.5	56.0	19938	6 ABP76679	Abp76679 Streptomy
44	49	53.3	64	4 AAU59596	Aau59596 Propionib
45	49	53.3	64	6 ABM56115	Abm56115 Propionib

ALIGNMENTS

RESULT 1

AAU78950
ID AAU78950 standard; peptide; 16 AA.

XX AC AAU78950;

XX 18-JUN-2002 (first entry)

XX Human PER1-protein transduction domain (PTD) fusion peptide.

XX Nuclear localisation signal; NLS; protein delivery; PER1; fusion protein;
KW membrane penetrating peptide; human; period 1;
KW protein transduction domain; membrane penetrating peptide.

XX Homo sapiens.

XX WO200218572-A2.

XX 07-MAR-2002.

XX 23-AUG-2001; 2001WO-US026421.

XX 25-AUG-2000; 2000US-0227647P.

XX 07-FEB-2001; 2001GB-00003110.

XX (AVET) AVENTIS PHARM INC.

XX Guo Y, Morse CC, Yao Z, Keesler GA;

XX WPI; 2002-304256/34.

XX New fusion proteins comprising membrane penetrating peptides, useful as
XX in vivo, ex vivo or in vitro intracellular carriers or delivery devices
XX for a compound of interest (e.g. peptide, protein, chemical entity,
XX nucleic acid).

XX Example 3; Page 29; 45pp; English.

XX This invention relates to a novel fusion protein, which comprises a
XX membrane penetrating peptide attached to a compound of interest. The
XX membrane penetrating peptide of the fusion protein is derived from a
XX nuclear localisation signal and may be the nuclear localisation signal
XX from human period protein hPER1. The fusion protein is useful for
XX delivery of a compound of interest into a cell. The fusion protein is
XX useful as in vivo, ex vivo or in vitro intracellular delivery devices for
XX a compound of interest (e.g. peptide, protein, chemical entity, nucleic
XX acid). In particular, the polypeptides are useful as protein carriers for
XX delivery of compounds to cells. The present sequence represents the human

firstly to send
per 15000

same invertebrate
assignee

period protein 1 (hPER1)-protein transduction domain (PTD) peptide used in an assay to analyse the ability of different mutant peptides to penetrate cellular membranes in the examples of the invention. This experiment was carried out to determine which amino acid residues of the hPER1 protein membrane penetrating peptide (MPP) are important for its function

XX SQ Sequence 16 AA;

Query Match 100.0%; Score 92; DB 5; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.1e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRHHCKSKAKRSRHH 16
Db 1 SRHHCKSKAKRSRHH 16

RESULT 2
AAU78934 ID AAU78934 standard; peptide; 16 AA.

XX AC AAU78934;

XX DT 18-JUN-2002 (first entry)

XX DE Human period protein 1 (hPER1) nuclear localisation signal.

XX KW Nuclear localisation signal; NLS; protein delivery; human; hPER1;
XX KW fusion protein; membrane penetrating peptide; period protein.

XX OS Homo sapiens.

XX FN WO200218572-A2.

XX PD 07-MAR-2002.

XX PF 23-AUG-2001; 2001WO-US026421.

XX PR 25-AUG-2000; 2000US-0227647P.

XX PR 07-FEB-2001; 2001GB-00003110.

XX PA (AVET) AVENTIS PHARM INC.

XX PI Guo Y, Morse CC, Yao Z, Keesler GA;

XX WPT; 2002-304256/34.

XX PT New fusion proteins comprising membrane penetrating peptides, useful as
XX PT in vivo, ex vivo or in vitro intracellular carriers or delivery devices
XX PT for a compound of interest (e.g. peptide, protein, chemical entity,
XX PT nucleic acid).

XX PS Example 2; Page 27; 45pp; English.

XX CC This invention relates to a novel fusion protein, which comprises a
XX CC membrane penetrating peptide attached to a compound of interest. The
XX CC membrane penetrating peptide of the fusion protein is derived from a
XX CC nuclear localisation signal and may be the nuclear localisation signal
XX CC from human period protein hPER1. The fusion protein is useful for
XX CC delivery of a compound of interest into a cell. The fusion protein is
XX CC useful as in vivo, ex vivo or in vitro intracellular delivery devices for
XX CC a compound of interest (e.g. peptide, protein, chemical entity, nucleic
XX CC acid). In particular, the peptide, protein, chemical entity, nucleic
XX CC delivery of compounds to cells. The present sequence represents the human
XX CC period protein hPER1 nuclear localisation signal (NLS) of the invention,
XX CC this NLS is rich in basic amino acids

XX SQ Sequence 16 AA;

Query Match 100.0%; Score 92; DB 5; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.1e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRHHCKSKAKRSRHH 16
Db 1 SRHHCKSKAKRSRHH 16

RESULT 3
ADB39032 ID ADB39032 standard; peptide; 16 AA.

XX AC ADB39032;

XX DT 04-DEC-2003 (first entry)

XX DE Human transcytosis peptide hPER1-1.

XX KW transduction sequence; targeting sequence; tumour associated peptide;
XX KW targeted immunogen; major histocompatibility complex; MHC;
XX KW tumour antigen; human melanoma antigen; gp100; MART-1; tyrosinase; MAGE;
XX KW TRP2; cytosolic; vaccine; anti-tumour; tumour; transcytosis peptide;
XX KW TAT; human period-1; human period-2; hPER1-1; hPER1-2; AntPHD;
XX KW Antennapedia homeodomain; transcytosis sequence; linker sequence;
XX KW epitope peptides; TAT-OVA peptide; hPER1-OVA peptide; hPER1-NP peptide;
XX KW hPER1-1-gp100; hPER1-2-gp100; AntPHD-gp100; human; transcytosis peptide;
XX KW hPER1-1.

XX OS Homo sapiens.

XX FN WO2003064609-A2.

XX PD 07-AUG-2003.

XX PF 29-JAN-2003; 2003WO-US002534.

XX PR 29-JAN-2002; 2002US-0352892P.

XX PR 15-AUG-2002; 2002US-00219850.

XX PA (AVET) AVENTIS PASTEUR LTD.

XX PA (AVET) AVENTIS PHARM INC.

XX PI Uger B, Salha D, Barber B, Morse B, Guo Y, Cheng S;

XX WPT; 2003-689527/65.

XX DR N-PSDB; ADB39064.

XX PT New immunogenic target polypeptides, useful for immunizing a host or
XX PT enhancing an anti-tumor immune response in a host, which protects the
XX PT host from the development of a tumor.

XX PS Claim 10; Page 33; 44pp; English.

XX CC This invention relates to a novel polypeptide consisting essentially of a
XX CC first amino acid sequence comprising a transduction (targeting) sequence
XX CC linked to a second amino acid sequence comprising a tumour associated
XX CC peptide. The invention also relates to methods for producing and
XX CC utilising targeted immunogens, preferably conjugating immunogens to an
XX CC amino acid sequence which targets the major histocompatibility complex
XX CC (MHC). The first amino acid sequence is derived from a tumour antigen,
XX CC preferably a human melanoma antigen such as gp100, MART-1, tyrosinase,
XX CC MAGE or TRP2. The polypeptides of the invention may have cytostatic
XX CC activity and may be of use in a vaccine. The polypeptides, DNA molecules
XX CC and compositions of the invention may therefore be useful for immunising
XX CC a host or enhancing an anti-tumour immune response in a host, which
XX CC protects the host from the development of a tumour. Transcytosis
XX CC peptides, such as TAT, human period-1 (hPER1)-1, hPER1-2 or AntPHD
XX CC (Antennapedia homeodomain), were selected for linking to the epitopes.
XX CC The epitope peptides were joined to the transcytosis sequence using a
XX CC linker sequence. The linker was selected from the sequence naturally
XX CC found directly N-terminal to the epitope sequence, or selected based on
XX CC known immunological parameters. Several immunogenic targets were
XX CC synthesised by combining the transcytosis peptides, linker sequences and
XX CC epitope peptides, such as TAT-OVA peptides, hPER1-OVA peptides, hPER1-NP
XX CC peptides, hPER1-1-gp100, hPER1-2-gp100 or AntPHD-gp100. The present
XX CC sequence is that of a transcytosis peptide which was used as a targeting

CC sequence for the polypeptides of the invention.

XX Sequence 16 AA;

Query Match 100.0%; Score 92; DB 7; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.1e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRRHCHSKAKSRHH 16
DB 1 SRRHCHSKAKSRHH 16

RESULT 4
ADE13829
ID ADE13829 standard; peptide; 16 AA.

XX

AC ADE13829;

XX 29-JAN-2004 (first entry)

XX Human period-1 protein (hPER1).

XX carcinoma embryonic antigen; CEA; CEA(6D)-1; 2; cytostatic; vaccine; cancer;

XX tumour antigen; immunotherapy.

XX Homo sapiens.

XX WO2003085087-A2.

XX 16-OCT-2003.

XX 09-APR-2003; 2003WO-US010916.

XX 09-APR-2002; 2002US-0372972P.

XX (AVET) AVENTIS PASTEUR LTD.

XX (THER-) THERION BIOLOGICS INC.

XX Parrington M, Zhang L, Rovinski B, Gritz LR, Greenhalgh T;

XX WPI; 2003-877029/81.

XX New isolated DNA molecule comprising the carcinoembryonic antigen (6D)-
1,2 sequence, useful for diagnosing, preventing and treating cancer, or
determining the effectiveness of a chemotherapeutic or other treatment
regimen.

XX Disclosure; Page 15; 56pp; English.

XX The present invention describes an isolated DNA molecule comprising the
carcinoembryonic antigen (CEA) (6D)-1,2 sequence of 2106 bp (see
ADE13829), or its fragment. Also described: (1) an expression vector
comprising the nucleic acid sequence CEA(6D)-1,2, or its fragment
described above; (2) a composition comprising the expression vector of
(1) in a pharmaceutical carrier; and (3) preventing or treating cancer
comprising administering to a host the expression vector of (1). CEA(6D)-
1,2 has cytostatic activity, and can be used in vaccines. The CEA(6D)-1,2
nucleic acid and target polypeptide are useful for diagnosing, preventing
and treating cancer, predicting prognosis, or determining the
effectiveness of a chemotherapeutic or other treatment regimen. The
expression vector may be used for the insertion and expression of CEA(6D)
-1,2 nucleic acid encoding tumour antigens for the immunotherapeutic
treatment of cancer. The target polypeptides are useful in generating
antibodies used in screening assays or for immunotherapy. The present
invention represents the human period-1 protein (hPER1), which is given in
the exemplification of the present invention.

XX Sequence 16 AA;

XX Query Match

XX Best Local Similarity 100.0%; Score 92; DB 7; Length 16;

XX Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRRHCHSKAKSRHH 16
DB 1 SRRHCHSKAKSRHH 16

RESULT 5

ADB39085
ID ADB39085 standard; peptide; 25 AA.

XX

AC ADB39085;

XX 04-DEC-2003 (first entry)

XX Human hPER1-1-gp100 (280-288) fusion peptide.

XX transduction sequence; targeting sequence; tumour associated peptide;
targeted immunogen; major histocompatibility complex; MHC;
tumour antigen; human melanoma antigen; gp100; MART-1; tyrosinase; MAGE;
TRP2; cytostatic; vaccine; anti-tumour; tumour; transcytosis peptide;
TAT; human period-1; human period-2; hPER1-1; hPER1-2; AntcPHD;

XX Antennapedia homeodomain; transcytosis sequence; linker sequence;
epitope peptides; TAT-OVA peptide; hPER1-OVA peptide; hPER1-NP peptide;
hPER1-1-gp100; hPER1-2-gp100; AntcPHD-gp100; human;
hPER1-1-gp100 (280-288) fusion peptide.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Misc-difference 1..16

XX /note= "This region is derived from hPER1"

XX Misc-difference 17..25

XX /note= "This region is derived from gp100"

XX WO2003084609-A2.

XX 07-AUG-2003.

XX 29-JAN-2003; 2003WO-US002534.

XX 29-JAN-2002; 2002US-0352892P.

XX 15-AUG-2002; 2002US-00219850.

XX (AVET) AVENTIS PASTEUR LTD.

XX (AVET) AVENTIS PHARM INC.

XX Uger B, Salha D, Barber B, Morse B, Guo Y, Cheng S;

XX WPI; 2003-689527/65.

XX New immunogenic target polypeptides, useful for immunizing a host or
enhancing an anti-tumour immune response in a host, which protects the
host from the development of a tumor.

XX Example 1; Page 24; 44pp; English.

XX This invention relates to a novel polypeptide consisting essentially of a
first amino acid sequence comprising a transduction (targeting) sequence
linked to a second amino acid sequence comprising a tumour associated
peptide. The invention also relates to methods for producing and
utilising targeted immunogens, preferably conjugating immunogens to an
amino acid sequence which targets the major histocompatibility complex
(MHC). The first amino acid sequence is derived from a tumour antigen,
preferably a human melanoma antigen such as gp100, MART-1, tyrosinase,
MAGE or TRP2. The polypeptides of the invention may have cytostatic
activity and may be of use in a vaccine. The polypeptides, DNA molecules
and compositions of the invention may therefore be useful for immunising
a host or enhancing an anti-tumour immune response in a host, which
protects the host from the development of a tumour. Transcytosis
peptides, such as TAT, human period-1 (hPER1)-1, hPER1-2 or AntcPHD
(Antennapedia homeodomain), were selected for linking to the epitopes.
The epitope peptides were joined to the transcytosis sequence using a
linker sequence. The linker was selected from the sequence naturally

CC found directly N-terminal to the epitope sequence, or selected based on
 CC known immunological parameters. Several immunogenic targets were
 CC synthesised by combining the transcytosis peptides, linker sequences and
 CC epitope peptides, such as TAT-OVA peptides, hPER1-OVA peptides, hPER1-NP
 CC peptides, hPER1-1-gp100, hPER1-2-gp100 or AntPHD-gp100. The present
 CC sequence is the amino acid sequence of a fusion peptide of the invention.
 XX
 XX

XX Sequence 25 AA;

Query Match 100.0%; Score 92; DB 7; Length 25;
 Best Local Similarity 100.0%; Pred. No. 1.3e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRRHCRSKAKRSRHH 16

Db 1 SRRHCRSKAKRSRHH 16

RESULT 6

ID ADB39082 standard; peptide; 25 AA.

AC ADB39082;

DT 04-DEC-2003 (first entry)

DE Human hPER1-NP fusion peptide 1.

XX transduction sequence; targeting sequence; tumour associated peptide;
 KW targeted immunogen; major histocompatibility complex; MHC;
 KW tumour antigen; human melanoma antigen; gp100; MART-1; tyrosinase; MAGE;
 KW TRP2; cytosolic; vaccine; anti-tumour; tumour; transcytosis peptide;
 KW TAT; human period-1; human period-2; hPER1-1; hPER1-2; AntPHD;
 KW Antennapedia homeodomain; transcytosis sequence; linker sequence;
 KW epitope peptides; TAT-OVA peptide; hPER1-OVA peptide; hPER1-NP peptide;
 KW hPER1-1-gp100; hPER1-2-gp100; AntPHD-gp100; human;
 KW hPER1-NP fusion peptide.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 1..16 /note= "This region is derived from hPER1"

FT Misc-difference 17..25

FT /note= "This region is derived from NP"

FT WO2003064609-A2.

XX 07-AUG-2003.

XX 29-JAN-2003; 2003WO-US002534.

XX 29-JAN-2002; 2002US-0352892P.

XX 15-AUG-2002; 2002US-00219850.

XX (AVET) AVENTIS PASTEUR LTD.

XX (AVET) AVENTIS PHARM INC.

XX Uger B, Salha D, Barber B, Morse B, Guo Y, Cheng S;

XX WPI; 2003-689527/65.

XX New immunogenic target polypeptides, useful for immunizing a host or
 PT enhancing an anti-tumour immune response in a host, which protects the
 PT host from the development of a tumor.

XX Example 1; Page 24; 44pp; English.

XX This invention relates to a novel polypeptide consisting essentially of a
 CC first amino acid sequence comprising a transduction (targeting) sequence
 CC linked to a second amino acid sequence comprising a tumour associated
 CC peptide. The invention also relates to methods for producing and
 CC utilising targeted immunogens, preferably conjugating immunogens to an

CC amino acid sequence which targets the major histocompatibility complex
 CC (MHC). The first amino acid sequence is derived from a tumour antigen,
 CC preferably a human melanoma antigen such as gp100, MART-1, tyrosinase,
 CC MAGE or TRP2. The polypeptides of the invention may have cytosolic
 CC activity and may be of use in a vaccine. The polypeptides, DNA molecules
 CC and compositions of the invention may therefore be useful for immunising
 CC a host or enhancing an anti-tumour immune response in a host, which
 CC protects the host from the development of a tumour. Transcytosis
 CC peptides, such as TAT, human period-1 (hPER1)-1, hPER1-2 or AntPHD
 CC (Antennapedia homeodomain), were selected for linking to the epitopes.
 CC The epitope peptides were joined to the transcytosis sequence using a
 CC linker sequence. The linker was selected from the sequence naturally
 CC found directly N-terminal to the epitope sequence, or selected based on
 CC known immunological parameters. Several immunogenic targets were
 CC synthesised by combining the transcytosis peptides, linker sequences and
 CC epitope peptides, such as TAT-OVA peptides, hPER1-OVA peptides, hPER1-NP
 CC peptides, hPER1-1-gp100, hPER1-2-gp100 or AntPHD-gp100. The present
 CC sequence is the amino acid sequence of a fusion peptide of the invention.
 XX
 XX

XX Sequence 25 AA;

Query Match 100.0%; Score 92; DB 7; Length 25;

Best Local Similarity 100.0%; Pred. No. 1.3e-07;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRRHCRSKAKRSRHH 16

Db 1 SRRHCRSKAKRSRHH 16

RESULT 7

AAU78964

ID AAU78964 standard; peptide; 40 AA.

XX AAU78964;

XX 18-JUN-2002 (first entry)

XX Human period 1 protein (hPER1) control peptide.

XX Nuclear localisation signal; NLS; protein delivery; PER1; fusion protein;
 KW membrane penetrating peptide; human; period 1;
 KW protein transduction domain; membrane penetrating peptide.

XX Homo sapiens.

XX WO200218572-A2.

XX 07-MAR-2002.

XX 23-AUG-2001; 2001WO-US026421.

XX 25-AUG-2000; 2000US-0227647P.

XX 07-FEB-2001; 2001GB-00003110.

XX (AVET) AVENTIS PHARM INC.

XX Guo Y, Morse CC, Yao Z, Keesler GA;

XX WPI; 2002-304256/34.

XX New fusion proteins comprising membrane penetrating peptides, useful as
 PT in vivo, ex vivo or in vitro intracellular carriers or delivery devices
 PT for a compound of interest (e.g. peptide, protein, chemical entity,
 PT nucleic acid).

XX Example 3; Page 30; 45pp; English.

XX This invention relates to a novel fusion protein, which comprises a
 CC membrane penetrating peptide attached to a compound of interest. The
 CC membrane penetrating peptide of the fusion protein is derived from a
 CC nuclear localisation signal and may be the nuclear localisation signal
 CC from human period protein hPER1. The fusion protein is useful for

CC delivery of a compound of interest into a cell. The fusion protein is
 CC useful as in vivo, ex vivo or in vitro intracellular delivery devices for
 CC a compound of interest (e.g. peptide, protein, chemical entity, nucleic
 CC acid). In particular, the polypeptides are useful as protein carriers for
 CC delivery of compounds to cells. The present sequence represents the human
 CC period protein 1 (PER1) hPER1 control peptide used in an assay to analyse
 CC the ability of different mutant peptides to penetrate cellular membranes
 CC in the examples of the invention. This experiment was carried out to
 CC determine which amino acid residues of the hPER1 protein membrane
 CC penetrating peptide (MPP) are important for its function
 XX Sequence 40 AA;

Query Match 100.0%; Score 92; DB 5; Length 40;
 Best Local Similarity 100.0%; Pred. No. 2.1e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRHHCRSKAKRSRHH 16
 DB 10 SRHHCRSKAKRSRHH 25

RESULT 8
 AAU78965
 ID AAU78965 standard; peptide; 70 AA.

XX AC AAU78965;
 DT 18-JUN-2002 (first entry)
 XX DE Human period 1 protein (hPER1) F7 fusion protein.
 XX KW Nuclear localisation signal; NLS; protein delivery; PER1; fusion protein;
 KW membrane penetrating peptide; human; period 1;
 KW protein transduction domain; membrane penetrating peptide.
 XX OS Homo sapiens.

XX FH Key Location/Qualifiers
 FT Misc-difference 35..65
 FT /label= Xaa
 FT /note= "Xaa= Unknown, these amino acids are represented
 FT by . . . in the specification"

XX PN WO200218572-A2.
 XX PD 07-MAR-2002.

XX PF 23-AUG-2001; 2001WO-US026421.

XX PR 25-AUG-2000; 2000US-0227647P.

XX PR 07-FEB-2001; 2001GB-00003110.

XX PA (AVET) AVENTIS PHARM INC.

XX PI Guo Y, Morse CC, Yao Z, Keesler GA;

XX XX WPI; 2002-304256/34.

XX New fusion proteins comprising membrane penetrating peptides, useful as
 FT in vivo, ex vivo or in vitro intracellular carriers or delivery devices
 FT for a compound of interest (e.g. peptide, protein, chemical entity,
 FT nucleic acid).

XX Example 1; Fig 1A; 45pp; English.

XX This invention relates to a novel fusion protein, which comprises a
 CC membrane penetrating peptide attached to a compound of interest. The
 CC membrane penetrating peptide of the fusion protein is derived from a
 CC nuclear localisation signal and may be the nuclear localisation signal
 CC from human period protein hPER1. The fusion protein is useful for
 CC delivery of a compound of interest into a cell. The fusion protein is
 CC useful as in vivo, ex vivo or in vitro intracellular delivery devices for

CC a compound of interest (e.g. peptide, protein, chemical entity, nucleic
 CC acid). In particular, the polypeptides are useful as protein carriers for
 CC delivery of compounds to cells. The present sequence represents the human
 CC period protein 1 (PER1) F7 fusion peptide of the invention
 XX Sequence 70 AA;

Query Match 100.0%; Score 92; DB 5; Length 70;
 Best Local Similarity 100.0%; Pred. No. 3.7e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRHHCRSKAKRSRHH 16
 DB 10 SRHHCRSKAKRSRHH 25

RESULT 9
 AAU06809
 ID AAU06809 standard; protein; 1290 AA.

XX AC AAU06809;
 DT 23-JUN-1999 (first entry)
 XX DE Human Per gene product.
 XX KW Circadian rhythm; Per gene; dimerisation domain; PAS-A; PAS-B; human;
 KW Drosophila; suprachiasmatic nucleus; SCN; brain; sleep stage regression;
 KW jet-lag; agrypnotic disorder; nightwalking; mouse.

XX OS Homo sapiens.

XX PN WO9914324-A1.

XX PD 25-MAR-1999.

XX PF 11-SEP-1998; 98WO-JP004125.

XX PR 12-SEP-1997; 97JP-00267846.

XX PA (SAKA/) SAKAKI Y.

XX PI Sakaki Y, Tei H;

XX DR WPI; 1999-229530/19.

XX DR N-PSDB; AAX32541.

XX Mammalian genes related to Drosophila circadian rhythm gene, used to
 FT treat, e.g. sleep regression.

XX PS Claim 4; Page 49-61; 76pp; Japanese.

XX The invention relates to mammalian (human and mouse) genes and their
 CC expression products which correspond to the Drosophila circadian rhythm
 CC gene dPer. The mammalian Per genes contain equivalents of the two
 CC dimerisation domains PAS-A and PAS-B in the Drosophila gene. Host cells
 CC transformed by vectors comprising the Per genes are used for the
 CC recombinant expression of their products. Per is expressed in the
 CC suprachiasmatic nucleus (SCN) of the brain and is involved in the
 CC regulation of the circadian rhythm which is largely controlled by that
 CC area. The mammalian Per genes and their products may be used for
 CC treatment and prevention of diseases in which the circadian rhythm is
 CC involved, such as sleep stage regression and progression syndromes, jet-
 CC lag and irregular or non-circadian agrypnotic disorders, and nightwalking
 CC by the demented. They also may be used in the occupational health
 CC management of persons working irregular night shifts. The present
 CC sequence represents a human Per gene product
 XX Sequence 1290 AA;

Query Match 100.0%; Score 92; DB 2; Length 1290;
 Best Local Similarity 100.0%; Pred. No. 7.7e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRHHCRSKAKRSRH 16
 DE - |||||
 DB 830 SRHHCRSKAKRSRH 845

RESULT 10
 ID AAY32216 standard; protein; 1290 AA.
 XX AAY32216;
 AC AAY32216;
 DT 15-FEB-2000 (first entry)
 XX Human PER protein.
 DE PER; period; perl gene; transcription factor; circadian rhythm; jet lag;
 KW sleep disorder; depression; seasonal affective disorder; fertility;
 KW therapy; human.
 XX Homo sapiens.
 OS WO9597137-A1.
 PN 11-NOV-1999.
 XX 06-MAY-1999; 99WO-US010072.
 PR 07-MAY-1998; 98US-0084610P.
 XX (HARD) HARVARD COLLEGE.
 PA Weitz CJ, Gekakis N, Staknis D;
 PI WPI; 2000-052938/04.
 DR N-PSDB; AA234630.
 XX Novel heterodimeric composition for identifying modulators used in
 PT diagnosing and treating circadian clock disruption disorders.
 XX Disclosure; Fig 28; 96pp; English.
 PS This sequence represents human PER protein. PER forms a heteromeric
 CC complex together with TIM (see AAY32220) protein and has a biological
 CC activity which inhibits transcription of the perl gene when the CLOCK
 CC protein (see AAY32215) is present in combination with BMAL1 protein (see
 CC AAY32209). The invention is based on the discovery of the transcriptional
 CC mechanism regulating genes responsible for the establishment and/or
 CC maintenance of the circadian clock, and provides an assay for novel
 CC modulators of BMAL1-CLOCK-mediated transcription of E-box-linked genes.
 CC The drugs are used to treat conditions such as jet lag, sleep disorders,
 CC depression (seasonal affective disorder) and infertility. The invention
 CC also provides BMAL1 and CLOCK proteins with which to stimulate the
 CC transcription of an E-box-linked gene which regulates the circadian clock
 XX Sequence 1290 AA;
 SQ

Query Match 100.0%; Score 92; DB 3; Length 1290;
 Best Local Similarity 100.0%; Pred. No. 7.7e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRHHCRSKAKRSRH 16
 DE - |||||
 DB 830 SRHHCRSKAKRSRH 845

RESULT 11
 ID ABB09289 standard; protein; 1290 AA.
 XX ABB09289;
 AC ABB09289;
 XX

DT 12-JUL-2002. (first entry)
 XX Human period (Drosophila) homologue 1 (PER1) protein SEQ ID NO.3.
 DE Human; period (Drosophila) homologue 1; PER1; polymorphic variant;
 KW polymorphic site; genotyping; haplotyping; circadian rhythm regulation;
 KW single nucleotide polymorphism; SNP.
 XX Homo sapiens.
 OS WO200222650-A2.
 PN 21-MAR-2002.
 XX 13-SEP-2001; 2001WO-US028780.
 PR 13-SEP-2000; 2000US-0232468P.
 XX (GENA-) GENAISSANCE PHARM INC.
 PA Duda A, Kliehm SE, Koshy B;
 PI WPI; 2002-393941/42.
 DR N-PSDB; ABL52078.
 XX Novel isolated human period Drosophila homolog 1 polynucleotide, useful
 PT for therapeutic purposes, for studying the expression and function of the
 PT polynucleotide, and for expressing the homolog.
 XX Claim 31; Fig 3; 162pp; English.
 PS The present invention describes an isolated human period (Drosophila)
 CC homologue 1, (PER1) polynucleotide (I) comprising a sequence which is a
 CC polymorphic variant for a reference sequence (ABL52077) for the PER1 gene
 CC or its fragment, or a polymorphic variant of a reference sequence also
 CC (ABL52078) for a PER1 cDNA or its fragment. The present invention also
 CC describes methods for genotyping and haplotyping the PER1 gene of an
 CC individual. (I) is useful in studying the expression and function of
 CC PER1, and in expressing PER1 protein for use in screening for candidate
 CC drugs to treat diseases related to PER1 activity. (I) is useful for
 CC therapeutic purposes. A recombinant non-human organism transformed or
 CC transfected with (I) can be used for studying expression of the PER1
 CC isogenes in vivo, for in vivo screening and testing of drugs targeted
 CC against PER1 protein, and for testing the efficacy of therapeutic agents
 CC and compounds for disorders associated with circadian rhythm regulation.
 CC The present sequence represents the human PER1 protein gene from the
 CC present invention
 XX Sequence 1290 AA;
 SQ

Query Match 100.0%; Score 92; DB 5; Length 1290;
 Best Local Similarity 100.0%; Pred. No. 7.7e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRHHCRSKAKRSRH 16
 DE - |||||
 DB 830 SRHHCRSKAKRSRH 845

RESULT 12
 ID AAY01687 standard; protein; 1291 AA.
 XX AAY01687;
 AC AAY01687;
 XX 23-JUN-1999 (first entry)
 DT Protein encoded by the 4.7 kb transcript of human RIGUI.
 DE RIGUI; Drosophila circadian rhythm period gene; circadian clock gene;
 KW Drosophila timeless ortholog.
 XX Homo sapiens.

XX WO9912952-A1.
 XX 18-MAR-1999.
 XX 09-SEP-1998; 98WO-US018755.
 XX 09-SEP-1997; 97US-0058256P.
 XX 04-NOV-1997; 97US-0065957P.
 XX (RERE-) RES DEV FOUND.
 XX Lee C, Albrecht U, Sun ZS, Eichele G;
 XX WPI; 1999-229221/19.
 XX N-PSDB; AAX26906.
 XX New isolated mammalian circadian rhythm genes.
 XX Claim 6; Fig 2; 73pp; English.
 XX The present sequence represents a RIGUI polypeptide. RIGUI is a gene
 CC corresponding to the Drosophila circadian rhythm period gene. The
 CC specification describes both mouse and human genes. The RIGUI
 CC polypeptides act as regulators of circadian rhythms. The identification
 CC of RIGUI as a putative circadian clock gene provides a useful tool to
 CC explore the molecular mechanism of the mammalian circadian machinery.
 CC Using interaction screening approaches, it should be possible to find
 CC interacting proteins, perhaps in the form of a Drosophila timeless
 CC ortholog. Furthermore, promoter analyses of the RIGUI gene should uncover
 CC how light cues and possibly other environmental stimuli, regulate the
 CC expression of this gene. Targeted disruption of the m-rigui gene using
 CC stem cell technology, may provide a valuable model system to study the
 CC various physiological and pathological aspects of disrupting circadian
 CC rhythms
 XX SQ Sequence 1291 AA;
 Query Match 100.0%; Score 92; DB 2; Length 1291;
 Best Local Similarity 100.0%; Pred. No. 7.7e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SRHHCRSKAKRSRHH 16
 |||||
 DB 830 SRHHCRSKAKRSRHH 845
 RESULT 13
 AAU78961
 ID AAU78961 standard; peptide; 16 AA.
 XX AC AAU78961;
 XX DT 18-JUN-2002 (first entry)
 XX DE Human PER1-PTD fusion peptide mutant R14A.
 XX KW Nuclear localisation signal; NLS; protein delivery; PER1; mutant;
 KW fusion protein; membrane penetrating peptide; human; period 1; mutein;
 KW protein transduction domain; membrane penetrating peptide.
 XX OS Homo sapiens.
 XX OS Synthetic.
 XX Key Location/Qualifiers
 FT Misc-difference 14 /note= "Wild type Arg replace by Ala".
 FT WO200218572-A2.
 XX PD 07-MAR-2002.
 XX PF 23-AUG-2001; 2001WO-US026421.

XX 25-AUG-2000; 2000US-0227647P.
 PR 07-FEB-2001; 2001GB-00003110.
 XX (AVET) AVENTIS PHARM INC.
 XX Guo Y, Morse CC, Yao Z, Keesler GA;
 XX WPI; 2002-304256/34.
 XX New fusion proteins comprising membrane penetrating peptides, useful as
 PT in vivo, ex vivo or in vitro intracellular carriers or delivery devices
 PT for a compound of interest (e.g. peptide, protein, chemical entity,
 PT nucleic acid).
 XX Example 3; Page 30; 45pp; English.
 XX This invention relates to a novel fusion protein, which comprises a
 CC membrane penetrating peptide attached to a compound of interest. The
 CC membrane penetrating peptide of the fusion protein is derived from a
 CC nuclear localisation signal and may be the nuclear localisation signal
 CC from human period protein hPER1. The fusion protein is useful for
 CC delivery of a compound of interest into a cell. The fusion protein is
 CC useful as in vivo, ex vivo or in vitro intracellular delivery devices for
 CC a compound of interest (e.g. peptide, protein, chemical entity, nucleic
 CC acid). In particular, the polypeptides are useful as protein carriers for
 CC delivery of compounds to cells. The present sequence represents the human
 CC period protein 1 (PER1)-protein transduction domain (PTD) peptide mutant
 CC R14A used in an assay to analyse the ability of different mutant peptides
 CC to penetrate cellular membranes in the examples of the invention. This
 CC experiment was carried out to determine which amino acid residues of the
 CC hPER1 protein membrane penetrating peptide (MPP) are important for its
 CC function
 XX SQ Sequence 16 AA;
 Query Match 96.7%; Score 89; DB 5; Length 16;
 Best Local Similarity 93.8%; Pred. No. 2.4e-07;
 Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SRHHCRSKAKRSRHH 16
 |||||
 DB 1 SRHHCRSKAKRSRHH 16
 RESULT 14
 AAU78957
 ID AAU78957 standard; peptide; 16 AA.
 XX AC AAU78957;
 XX DT 18-JUN-2002 (first entry)
 XX DE Human PER1-PTD fusion peptide mutant S8A.
 XX KW Nuclear localisation signal; NLS; protein delivery; PER1; mutant;
 KW fusion protein; membrane penetrating peptide; human; period 1; mutein;
 KW protein transduction domain; membrane penetrating peptide.
 XX OS Homo sapiens.
 XX OS Synthetic.
 XX Key Location/Qualifiers
 FT Misc-difference 8 /note= "Wild type Ser replace by Ala"
 FT WO200218572-A2.
 XX PD 07-MAR-2002.
 XX PF 23-AUG-2001; 2001WO-US026421.
 XX PF 25-AUG-2000; 2000US-0227647P.

```
PR 07-FEB-2001; 2001GB-00003110.
XX
PA (AVET-) AVENTIS PHARM INC.
XX
PI Guo Y, Morse CC, Yao Z, Keesler GA;
XX
DR WPI; 2002-304256/34.
XX
XX New fusion proteins comprising membrane penetrating peptides, useful as
PT in vivo, ex vivo or in vitro intracellular carriers or delivery devices
PT for a compound of interest (e.g. peptide, protein, chemical entity,
PT nucleic acid).
XX
PS Example 3; Page 30; 45pp; English.
XX
XX This invention relates to a novel fusion protein, which comprises a
CC membrane penetrating peptide attached to a compound of interest. The
CC membrane penetrating peptide of the fusion protein is derived from a
CC nuclear localisation signal and may be the nuclear localisation signal
CC from human period protein hPER1. The fusion protein is useful for
CC delivery of a compound of interest into a cell. The fusion protein is
CC useful as in vivo, ex vivo or in vitro intracellular delivery devices for
CC a compound of interest (e.g. peptide, protein, chemical entity, nucleic
CC acid). In particular, the polypeptides are useful as protein carriers for
CC delivery of compounds to cells. The present sequence represents the human
CC period protein 1 (PER1)-protein transduction domain (PTD) peptide mutant
CC S8A used in an assay to analyse the ability of different mutant peptides
CC to penetrate cellular membranes in the examples of the invention. This
CC experiment was carried out to determine which amino acid residues of the
CC hPER1 protein membrane penetrating peptide (MPP) are important for its
CC function.
XX
SQ Sequence 16 AA;
Query Match 96.7%; Score 89; DB 5; Length 16;
Best Local Similarity 93.8%; Pred. No. 2.4e-07;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SRHHCRSKAKRSRHH 16
DB 1 SRHHCRSKAKRSRHH 16
RESULT 15
AAU78926
ID AAU78926 standard; peptide; 16 AA.
XX
AC AAU78926;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human period protein 1 (hPER1) peptide.
XX
KW Nuclear localisation signal; NLS; protein delivery; human; hPER1;
KW fusion protein; membrane penetrating peptide; human period protein.
XX
OS Homo sapiens.
XX
FN WO200218572-A2.
XX
PD 07-MAR-2002.
XX
PF 23-AUG-2001; 2001WO-US026421.
XX
PR 25-AUG-2000; 2000US-0227647P.
PR 07-FEB-2001; 2001GB-00003110.
XX
XX (AVET ) AVENTIS PHARM INC.
XX
PI Guo Y, Morse CC, Yao Z, Keesler GA;
XX
XX WPI; 2002-304256/34.
XX
XX New fusion proteins comprising membrane penetrating peptides, useful as
PT in vivo, ex vivo or in vitro intracellular carriers or delivery devices
PT for a compound of interest (e.g. peptide, protein, chemical entity,
PT nucleic acid).
XX
PS Example 2; Page 27; 45pp; English.
XX
XX This invention relates to a novel fusion protein, which comprises a
CC membrane penetrating peptide attached to a compound of interest. The
CC membrane penetrating peptide of the fusion protein is derived from a
CC nuclear localisation signal and may be the nuclear localisation signal
CC from human period protein hPER1. The fusion protein is useful for
CC delivery of a compound of interest into a cell. The fusion protein is
CC useful as in vivo, ex vivo or in vitro intracellular delivery devices for
CC a compound of interest (e.g. peptide, protein, chemical entity, nucleic
CC acid). In particular, the polypeptides are useful as protein carriers for
CC delivery of compounds to cells. The present sequence represents the human
CC period protein 1 (hPER1) used in an assay to analyse the ability of
CC different peptides to penetrate cellular membranes in the examples of the
CC invention.
XX
SQ Sequence 16 AA;
Query Match 95.7%; Score 88; DB 5; Length 16;
Best Local Similarity 100.0%; Pred. No. 3.5e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 RRHHCRSKAKRSRHH 16
DB 2 RRHHCRSKAKRSRHH 16
Search completed: May 21, 2004, 12:53:27
Job time : 56 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2004, 12:50:55 ; Search time 21 Seconds
(without alignments)
73.289 Million cell updates/sec

Title: US-09-933-780C-16
Perfect score: 92
Sequence: 1 SRRHCRSKAKRSRHH 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92	100.0	1290	T00018	period protein hom
2	88	95.7	1291	T00019	period protein hom
3	49	53.3	325	T09667	peroxidase (EC 1.1
4	49	53.3	517	A54099	protein kinase Dar
5	48	52.2	611	EDBSM	immediate-early pr
6	47	51.1	375	F69551	coenzyme POQ synth
7	45.5	49.5	288	S68798	RNA-binding protei
8	45	48.9	135	E96750	hypothetical prote
9	45	48.9	503	JW0046	estrogen receptor
10	44	47.8	364	D97549	hypothetical prote
11	43	46.7	163	T28012	hypothetical prote
12	43	46.7	197	T10413	hypothetical prote
13	43	46.7	247	T45847	hypothetical prote
14	43	46.7	302	JQ2275	serotonin receptor
15	43	46.7	477	S71400	estrogen receptor
16	43	46.7	530	JC5939	estrogen receptor
17	43	46.7	1235	T17457	SARA protein - Afr
18	42.5	46.2	56	C58213	protamine II - Ame
19	42	45.7	102	S33336	protamine P2 - rhe
20	42	45.7	102	S33335	protamine P2 - com
21	42	45.7	103	S33337	protamine P2 - pig
22	42	45.7	113	S66936	probable membrane
23	42	45.7	622	S45129	VPS27 protein - ye
24	42	45.7	1208	B82091	exodeoxyribonuclea
25	42	45.7	1280	T42514	kinase anchor prot
26	41.5	45.1	219	S28507	transcription fact
27	41.5	45.1	1038	T13177	sog protein - frui
28	41	44.6	107	A29995	protamine P2 precu
29	41	44.6	254	S40944	hypothetical prote

30	41	44.6	266	2	D88567	protein ZK632.12 l
31	41	44.6	301	2	S76240	hypothetical prote
32	41	44.6	316	2	T13601	hypothetical prote
33	41	44.6	385	2	T18821	hypothetical prote
34	41	44.6	454	2	A13467	glycine betaine/l-
35	41	44.6	530	2	E89044	protein B0238.9 li
36	41	44.6	575	2	T34509	hypothetical prote
37	41	44.6	658	2	T33568	hypothetical prote
38	41	44.6	926	2	T15683	hypothetical prote
39	41	44.6	967	2	H86334	T20H2.10 protein -
40	41	44.6	1012	2	A84393	glycolate oxidase
41	41	44.6	1016	2	T00375	hypothetical prote
42	41	44.6	1107	2	T21280	hypothetical prote
43	41	44.6	1757	2	T05204	hypothetical prote
44	41	44.6	2241	2	T02857	conserved hypothet
45	40	43.5	102	2	S33334	protamine P2 - ora

ALIGNMENTS

RESULT 1

T00018
period protein homolog - human
C:Species: Homo sapiens (man)
C>Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
C:Accession: T00018
R:Tei, H.; Okamura, H.; Shigeyoshi, Y.; Fukuhara, C.; Ozawa, R.; Hirose, M.; Sakaki, Y
Nature 389, 512-516, 1997
A>Title: Circadian oscillation of a mammalian homologue of the Drosophila period gene.
A:Reference number: Z14056; MUID:97472418; PMID:9333243
A:Accession: T00018
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1290 <TEI>
A:Cross-references: EMBL:AB002107; NID:g2506044; PIDN:BA22633.1; PID:g2506045
A:Experimental source: brain
C:Genetics:
A:Gene: hper
A:Map position: 17q12-13.1

Query Match 100.0%; Score 92; DB 2; Length 1290;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRRHCRSKAKRSRHH 16
|||
Db 830 SRRHCRSKAKRSRHH 845

RESULT 2

T00019
period protein homolog - mouse
C:Species: Mus musculus (house mouse)
C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 21-Jul-2000
C:Accession: T00019
R:Tei, H.; Okamura, H.; Shigeyoshi, Y.; Fukuhara, C.; Ozawa, R.; Hirose, M.; Sakaki, Y
Nature 389, 512-516, 1997
A>Title: Circadian oscillation of a mammalian homologue of the Drosophila period gene.
A:Reference number: Z14056; MUID:97472418; PMID:9333243
A:Accession: T00019
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1291 <TEI>
A:Cross-references: EMBL:AB002108; NID:g2506046; PIDN:BA22634.1; PID:g2506047
A:Experimental source: adult brain
C:Genetics:
A:Gene: mper
A:Map position: 11B

Query Match 95.7%; Score 88; DB 2; Length 1291;
Best Local Similarity 100.0%; Pred. No. 6.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRHCRSKAKRSRHH 16
 |||||
 DB 828 RRHCRSKAKRSRHH 842

RESULT 3

T09667
 C:Species: EC 1.11.1.7) pxdd precursor - alfalfa (fragment)
 C:Species: Medicago sativa (alfalfa)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 26-Aug-1999
 C:Accession: T09667
 R:Abraham, S.L.; Hayes, C.M.; Watson, J.M.
 A:Description: Organ-specific expression of three peroxidase-encoding cDNAs from lucerne submitted to the EMBL Data Library, September 1994
 A:Reference number: Z16809
 A:Accession: T09667
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-325 <ABR>
 A:Cross-references: EMBL:L36158; NID:G537318; PID:G537319
 A:Experimental source: subspecies sativa; cultivar Siriver
 C:Genetics:
 A:Note: pxdd
 C:Superfamily: peroxidase
 C:Keywords: heme; iron; oxidoreductase
 F:1-22/Domain: signal sequence (fragment) #status predicted <SIG>
 F:23-325/Product: peroxidase pxdd #status predicted <MAT>

Query Match 53.3%; Score 49; DB 2; Length 325;
 Best Local Similarity 61.5%; Pred. No. 2.5;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 HCRSKAKRSRHH 16
 |||||
 DB 16 HSCRTHAQLSRHH 28

RESULT 4

A54099
 Protein kinase Darkener-of-apricot (EC 2.7.1.1) - fruit fly (Drosophila melanogaster)
 N:Alternate names: LAMMER protein kinase Doa
 C:Species: Drosophila melanogaster
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 03-May-2002
 C:Accession: A54099; S44077
 R:Yun, B.; Parkas, R.; Lee, K.; Rabinow, L.
 Genes Dev. 8, 1160-1173, 1994
 A:Title: The Doa locus encodes a member of a new protein kinase family and is essential
 A:Reference number: A54099; MUID:95011531; PMID:7926721
 A:Accession: A54099
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-517 <YUN>
 A:Cross-references: GB:X87815; NID:G472912; PIDN:CAA55367.1; PID:G472913
 C:Genetics:
 A:Gene: FlyBase:Doa
 A:Cross-references: FlyBase:FBgn0000480
 C:Superfamily: human protein kinase ckl1; protein kinase homology
 C:Keywords: ATP; autophosphorylation; phosphotransferase; protein kinase
 F:168-438/Domain: protein kinase homology <KIN>

Query Match 53.3%; Score 49; DB 2; Length 517;
 Best Local Similarity 75.0%; Pred. No. 3.6;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 HCRSKAKRSRHH 15
 |||||
 DB 53 HHTSAKRRHH 64

RESULT 5

EDBESM
 immediate-early protein IE3 - murine cytomegalovirus (strain Smith)

C:Species: murine cytomegalovirus, murine herpesvirus 1
 C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 08-Apr-1994
 C:Accession: A40835
 R:Messerle, M.; Buehler, B.; Keil, G.M.; Koszinowski, U.H.
 J. Virol. 66, 27-36, 1992
 A:Title: Structural organization, expression, and functional characterization of the mt
 A:Reference number: A40835; MUID:92085392; PMID:1309246
 A:Accession: A40835
 A:Molecule type: mRNA
 A:Residues: 1-611 <MES>
 A:Cross-references: GB:M77846
 C:Superfamily: murine cytomegalovirus immediate-early phosphoprotein pp89
 C:Keywords: immediate-early protein

Query Match 52.2%; Score 48; DB 1; Length 611;
 Best Local Similarity 56.2%; Pred. No. 6;
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 SRRHCRSKAKRSRHH 16
 |||||
 DB 277 SRRHCRSKAKRSRHH 292

RESULT 6

F69551
 coenzyme PQQ synthesis protein (pqqE) homolog - Archaeoglobus fulgidus
 C:Species: Archaeoglobus fulgidus
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
 C:Accession: F69551
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
 A:Reference number: A69250; MUID:98049343; PMID:9389475
 A:Accession: F69551
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-375 <KLE>
 A:Cross-references: GB:AE001109; GB:AE000782; NID:G2689432; PIDN:AAB91253.1; PID:G26506

Query Match 51.1%; Score 47; DB 2; Length 375;
 Best Local Similarity 72.7%; Pred. No. 5.7;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 HCRSKAKRSRHH 15
 |||||
 DB 21 HCRAKAIRKHH 31

RESULT 7

S68798
 RNA-binding protein SIG41 - mouse
 N:Alternate names: Tra2 splicing factor homolog
 C:Species: Mus musculus (house mouse)
 C:Date: 25-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 20-Jun-2000
 C:Accession: S68798
 R:Segade, F.; Hurlé, B.; Claudio, E.; Ramos, S.; Lazo, P.S.
 FEBS Lett. 387, 152-156, 1996
 A:Title: Molecular cloning of a mouse homologue for the Drosophila splicing regulator 1
 A:Reference number: S68798; MUID:96244514; PMID:8674539
 A:Accession: S68798
 A:Molecule type: mRNA
 A:Residues: 1-288 <SEG>
 A:Cross-references: GB:X80232; NID:G1279557; PIDN:CAA56518.1; PID:G1279558
 C:Genetics:
 A:Gene: SIG41

C:Superfamily: transformer-2 sex-determining protein; ribonucleoprotein repeat homology
 F:119-186/Domain: ribonucleoprotein repeat homology <RRN>
 F:120-125/Region: RNA-binding RNP2 motif
 F:159-166/Region: RNA-binding RNP1 motif

F;242-249/Region: glycine-rich

Query Match 49.5%; Score 45.5; DB 2; Length 288;
Best Local Similarity 62.5%; Pred. No. 7.9;
Matches 10; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 1 SRHHCKRSKAKRRHH 16
||||:|::|||
Db 64 SRHHYTSRS-RSRSH 78

RESULT 8
E96750
hypothetical protein F28P22.21 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: E96750
A:;Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Li, J.H.; Lin, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, C.W.; Rooney, T.; Rowley, D.; Sakano, H.
A:;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: E96750
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-135 <STO>
A:Cross-references: GB:AE005173; NID:96648168; PIDN:AAF21168.1; GSPDB:GN00141
C:Genetics:
A:Gene: F28P22.21
A:Map position: 1

Query Match	48.9%;	Score 45;	DB 2;	Length 135;
Best Local Similarity	46.7%;	Pred. No. 5;		
Matches	7;	Conservative	3;	Mismatches 5; Indels 0; Gaps 0;
QY	2	RRHHCSKAKRSPHH	16	
Db	68	KHHHREKKWRQRHH	82	

RESULT 9
JW0046
estrogen receptor beta2 - rat
N:Alternate names: ERbeta2
C:Species: Rattus norvegicus (Norway rat)
C:Date: 17-Jun-1998 #sequence_revision 10-Jul-1998 #text_change 17-Mar-1999
C:Accession: JW0046
R:R.Mariyana, K.; Endoh, H.; Sasaki-Iwaka, H.; Kanou, H.; Shimaya, E.; Hashimoto, S.; Kat
A:Title: A novel isoform of rat estrogen receptor beta with 18 amino acid insertion in t
A:Reference number: JW0046; MUID:98262932; PMID:960083
A:Accession: JW0046
A:Molecule type: mRNA
A:Residues: 1-503. <VAR>
A:Cross-references: DDBJ:AB012721
C:Comment: This protein functions as a negative regulator of estrogen action.
C:Superfamily: estrogen receptor; ERBA transforming protein homology
F:102-381/Domain: ERBA transforming protein homology <ERBA>

```

Query Match      48.9%;   Score 45;   DB 2;   Length 503;
Best Local Similarity 56.2%;   Pred. NO. 15;
Matches 9;   Conservative 2;   Mismatches 5;   Indels 0;   Gaps 0;

Qy      1  SRRHHCRSKAKGSRHH  16
      | : | | | | | |
Db      191 SEOVHCLSKAKNGCH  206

```

RESULT 10

D97549
hypothetical protein AGR_C_2880 [imported] - Agrobacterium tumefaciens (strain C58, C58, C58)
C:Species: Agrobacterium tumefaciens
C:date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C:Accession: D97549
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurolo, B.; Goldman, B.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, E.
S:Science 294, 2323-2328, 2001
A:title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium t
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: D97549
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-364 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK87349.1; PID:gl5156651; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_2880
A:Map position: circular chromosome

Query Match	47.8%;	Score 44;	DB 2;	Length 364;
Best Local Similarity	53.3%;	Pred. No. 16;		
Matches	8;	Conservative	2;	Mismatches 5; Indels 0; Gaps 0;
QY	2	RRHHCRSKAKRRSRHH	16	
		: : : :		
Db	53	RKHHPGAKQSFH	67	

RESULT 11

ZK813.1
 hypothetical protein ZK813.1 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
 C:Accession: T28012
 R:leimbac, D.
 submitted to the EMBL Data Library, November 1995
 A:Description: The sequence of C. elegans cosmid ZK813.
 A:Reference number: Z20455
 A:Accession: T28012
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-163 <LET>
 A:Cross-references: EMBL:U40954; PIDN:AA852654.1; GSPDB:GN00028; CESP:ZK813.1
 A:Experimental source: strain Bristol N2; clone ZK813
 C:Genetics: .
 A:Gene: CESP:ZK813.1
 A:Map position: X
 A:Introns: 36/2; 145/2

RESULT 12

T10413
Hypothetical protein 144 - *Orygia pseudotsugata* nuclear polyhedrosis virus
C:Species: *Orygia pseudotsugata* nuclear polyhedrosis virus, OpMPVP
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999.#text_change 21-Jul-2000
C:Accession: T10413
R:Aluhrs, C.J.; Russell, R.R.; Funk, C.J.; Evans, J.; Harwood, S.; Rohmann, G.F.
Virology 229, 381-399, 1997
A:Title: The sequence of the *Orygia pseudotsugata* multinucleocapsid nuclear polyhedrosis
A:Reference number: Z17011; MUID:97271300; PMID:9126251
A:Accession: T10413
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

A:Residues: 1-197 <AHR>

A:Cross-references: EMBL:U75930; NID:G2934903; PIDN:AAC59143.1; PID:GI911390

Query Match 46.7%; Score 43; DB 2; Length 197;

Best Local Similarity 72.7%; Pred. No. 14;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY * 5 HCRSKAKRSRH 15

Db 113 HRRSEAKRTH 123

RESULT 13

T45847

hypothetical protein F3A4.20 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 28-Jul-2000

C:Accession: T45847

R:Barques, M.; Collado, M.C.; Navarro, P.; Terol, J.; Perez-Alonso, M.; Mewes, H.W.; May

submitted to the Protein Sequence Database, December 1999

A:Reference number: 223007

A:Accession: T45847

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-247 <BAR>

A:Cross-references: EMBL:AL132978

A:Experimental source: cultivar Columbia; BAC clone F3A4

C:Genetics:

A:Map position: 3

A:Introns: 62/1

A:Note: F3A4.20

C:Superfamily: Arabidopsis thaliana hypothetical protein F3A4.20

Query Match 46.7%; Score 43; DB 2; Length 247;

Best Local Similarity 63.8%; Pred. No. 17;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 HCRSKAKRSR 14

Db 153 HCRPSSRSR 163

RESULT 14

JQ2275

serotonin receptor 1D pseudogene - human

N:Alternate names: 5-hydroxytryptamine receptor 1D pseudogene (5-HTR1Dpsi)

C:Species: Homo sapiens (man)

C>Date: 30-Sep-1993 #sequence_revision 09-May-1996 #text_change 20-Apr-2000

C:Accession: JQ2275

R:Nguyen, T.; Marchese, A.; Kennedy, J.L.; Petronis, A.; Peroutka, S.J.; Wu, P.H.; O'Dow

Gene 124, 295-301, 1993

A:Title: An Alu sequence interrupts a human 5-hydroxytryptamine 1D receptor pseudogene.

A:Reference number: JQ2275; MUID:93185939; PMID:8444354

A:Accession: JQ2275

A>Status: conceptual translation of pseudogene

A:Molecule type: DNA

A:Residues: 1-302 <NGU>

A:Cross-references: GB:106179; NID:G184436

A:Note: this sequence represents an approximate translation of a pseudogene, constructed

ile the original reading frame is preserved

C:Genetics:

A:Gene: GDB:HTR1DP1

A:Cross-references: GDB:138784

A:Map position: 12pter-12qter

C:Keywords: pseudogene

Query Match 46.7%; Score 43; DB 4; Length 302;

Best Local Similarity 50.0%; Pred. No. 20;

Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 HCRSKAKRSRH 16

Db 177 HCPTRTSRSHH 188

RESULT 15

S71400

estrogen receptor beta - human

C:Species: Homo sapiens (man)

C>Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 20-Jun-2000

C:Accession: S71400

R:Mosselman, S.; Polman, J.; Dijkema, R.

FEBS Lett. 392, 49-53, 1996

A:Title: ER-beta: identification and characterization of a novel human estrogen receptor

A:Reference number: S71400; MUID:96354875; PMID:8769313

A:Accession: S71400

A:Molecule type: mRNA

A:Residues: 1-477 <MOS>

A:Cross-references: EMBL:X99101; NID:G1518262; PIDN:CAA67555.1; PID:G1518263

C:Superfamily: estrogen receptor; erba transforming protein homology

C:Keywords: DNA binding; nucleus; phosphoprotein; steroid binding; steroid hormone rec

F:94-355/Domain: erba transforming protein homology <ERBA>

F:96-116/Region: zinc finger CCCC motif

F:132-156/Region: zinc finger CCCC motif

F:167-182/Region: nuclear location signal

F:52/Binding site: phosphate (Ser) (covalent) (by MAP kinase) #status predicted

F:96-99,113,116/Binding site: zinc (Cys) #status predicted

F:132,138,148,151/Binding site: zinc (Cys) #status predicted

Query Match

Best Local Similarity 66.7%; Score 43; DB 2; Length 477;

Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 HCRSKAKRSRH 16

Db 187 HCAGKAKRSRGGH 198

Search completed: May 21, 2004, 12:55:17

Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2004, 12:52:25 ; Search time 23 Seconds
(without alignments)
35.914 Million cell updates/sec

Title: US-09-933-780C-16

Perfect score: 92

Sequence: 1 SRHHCRSKAKRSRHH 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA:*

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/2/iaa/PTCUS_COMB.pep:*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92	100.0	1290	3	US-09-150-460B-6
2	88	95.7	1291	3	US-09-150-460B-10
3	88	95.7	1291	3	US-09-220-641-5
4	49	53.3	517	4	US-09-457-040B-14
5	48	52.2	394	4	US-09-489-039A-8445
6	47.5	51.6	357	4	US-09-252-991A-20112
7	47	51.1	23	4	US-09-693-822B-2
8	47	51.1	192	4	US-09-252-991A-28554
9	47	51.1	618	4	US-09-252-991A-23696
10	47	51.1	618	4	US-09-252-991A-28358
11	46	50.0	19	4	US-09-693-822B-18
12	46	50.0	177	4	US-09-252-991A-16889
13	45	48.9	248	4	US-09-252-991A-27856
14	45	48.9	422	4	US-09-724-864-45
15	45	48.9	484	2	US-08-836-620A-13
16	-45	48.9	485	2	US-08-836-620A-2
17	45	48.9	574	4	US-09-252-991A-31651
18	44	47.8	15	4	US-09-693-822B-22
19	44	47.8	16	4	US-09-693-822B-21
20	44	47.8	17	4	US-09-693-822B-20
21	44	47.8	18	4	US-09-693-822B-19
22	44	47.8	19	4	US-09-693-822B-3
23	44	47.8	19	4	US-09-693-822B-4
24	44	47.8	19	4	US-09-693-822B-5
25	44	47.8	19	4	US-09-693-822B-6
26	44	47.8	19	4	US-09-693-822B-7
27	44	47.8	19	4	US-09-693-822B-8

28	44	47.8	19	4	US-09-693-822B-23	Sequence 23, Appl
29	44	47.8	19	4	US-09-693-822B-24	Sequence 24, Appl
30	44	47.8	19	4	US-09-693-822B-26	Sequence 26, Appl
31	44	47.8	700	4	US-09-252-991A-28344	Sequence 28344, A
32	43	46.7	19	4	US-09-693-822B-25	Sequence 25, Appl
33	43	46.7	384	2	US-08-836-620A-15	Sequence 15, Appl
34	43	46.7	412	4	US-09-252-991A-19536	Sequence 19536, A
35	43	46.7	416	4	US-09-608-088-6	Sequence 6, Appl
36	43	46.7	418	4	US-09-608-088-21	Sequence 21, Appl
37	43	46.7	477	4	US-09-608-088-5	Sequence 5, Appl
38	43	46.7	484	2	US-08-836-620A-14	Sequence 14, Appl
39	43	46.7	485	2	US-08-836-620A-3	Sequence 3, Appl
40	43	46.7	485	2	US-08-836-620A-5	Sequence 5, Appl
41	43	46.7	522	4	US-09-252-991A-26377	Sequence 26377, A
42	43	46.7	530	4	US-09-608-088-25	Sequence 25, Appl
43	43	46.7	548	3	US-09-139-617-1	Sequence 1, Appl
44	43	46.7	548	4	US-09-561-741A-1	Sequence 1, Appl
45	43	46.7	548	4	US-09-558-795-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1

US-09-150-460B-6

; Sequence 6, Application US/09150460B

; Patent No. 6190882

; GENERAL INFORMATION:

; APPLICANT: Lee, Cheng-Chi

; APPLICANT: Albrecht, Urs

; APPLICANT: Eichele, Gregor

; APPLICANT: Sun, Zhong Sheng

; TITLE OF INVENTION: Mammalian Circadian Rhythm-Like Gene

; FILE REFERENCE: D6039

; CURRENT APPLICATION NUMBER: US/09/150,460B

; CURRENT FILING DATE: 1998-09-09

; PRIOR APPLICATION NUMBER: US 60/058,256

; PRIOR FILING DATE: 1997-09-09

; NUMBER OF SEQ ID NOS: 21

; SEQ ID NO 6

; LENGTH: 1290

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: Protein sequence corresponding to RIGUI 4.7; Gene Bank

; OTHER INFORMATION: Accession Number: AF022991

US-09-150-460B-6

Query Match

Best Local Similarity 100.0%; Score 92; DB 3; Length 1290;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRHHCRSKAKRSRHH 16

DB 830 SRHHCRSKAKRSRHH 845

RESULT 2

US-09-150-460B-10

; Sequence 10, Application US/09150460B

; Patent No. 6190882

; GENERAL INFORMATION:

; APPLICANT: Lee, Cheng-Chi

; APPLICANT: Albrecht, Urs

; APPLICANT: Eichele, Gregor

; APPLICANT: Sun, Zhong Sheng

; TITLE OF INVENTION: Mammalian Circadian Rhythm-Like Gene

; FILE REFERENCE: D6039

; CURRENT APPLICATION NUMBER: US/09/150,460B

; CURRENT FILING DATE: 1998-09-09

; PRIOR APPLICATION NUMBER: US 60/058,256

; PRIOR FILING DATE: 1997-09-09

; NUMBER OF SEQ ID NOS: 21

; SEQ ID NO 10
; LENGTH: 1291
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: Protein sequence encoded by m-rigui homologue
US-09-150-460B-10

Query Match 95.7%; Score 88; DB 3; Length 1291;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RHHCERSKAKRSRH 16
Db 828 RHHCERSKAKRSRH 842

RESULT 3
US-09-220-641-5
; Sequence 5, Application US/09220641
; Patent No. 6210923
; GENERAL INFORMATION:
; APPLICANT: Lee, Cheng-Chi
; APPLICANT: Sun, Zhong Sheng
; APPLICANT: Albrecht, Urs
; APPLICANT: Eichele, Gregor
; TITLE OF INVENTION: Mammalian Circadian Regulator M-RIGUI2 (M-PER2)
; FILE REFERENCE: D6067
; CURRENT APPLICATION NUMBER: US/09/220,641
; CURRENT FILING DATE: 1998-12-24
; EARLIER APPLICATION NUMBER: US 60/068,886
; EARLIER FILING DATE: 1997-12-26
; NUMBER OF SEQ ID NOS: 5
; SEQ ID NO 5
; LENGTH: 1291
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: m-rigui1
US-09-220-641-5

Query Match 95.7%; Score 88; DB 3; Length 1291;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RHHCERSKAKRSRH 16
Db 828 RHHCERSKAKRSRH 842

RESULT 4
US-09-457-040B-14
; Sequence 14, Application US/09457040B
; Patent No. 6387641
; GENERAL INFORMATION:
; APPLICANT: Vertex Pharmaceuticals Incorporated
; APPLICANT: Bellon, Steve
; TITLE OF INVENTION: Crystallized P38 Complexes
; FILE REFERENCE: VPI/98-14
; CURRENT APPLICATION NUMBER: US/09/457,040B
; CURRENT FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 517
; TYPE: PRT
; ORGANISM: DROME - Drosophila Melanogaster Fruit Fly
US-09-457-040B-14

Query Match 53.3%; Score 49; DB 4; Length 517;
Best Local Similarity 75.0%; Pred. No. 7.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 HHCRSKAKRSRH 15
Db 53 HHTESAARRRH 64

RESULT 5
US-09-489-039A-8445
; Sequence 8445, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8445
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8445

Query Match 52.2%; Score 48; DB 4; Length 394;
Best Local Similarity 66.7%; Pred. No. 7.8;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 RHHCERSKAKRS 13
Db 84 RHHLRAKARRA 95

RESULT 6
US-09-252-991A-20112
; Sequence 20112, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20112
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20112

Query Match 51.6%; Score 47.5; DB 4; Length 357;
Best Local Similarity 71.4%; Pred. No. 8.4;
Matches 10; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy 1 SRRHH-CRSKAKRS 13
Db 82 TRHHSCSLARRS 95

RESULT 7
US-09-693-822B-2
; Sequence 2, Application US/09693822B
; Patent No. 6555650
; GENERAL INFORMATION:
; APPLICANT: LaJoie, Gilles A.
; TITLE OF INVENTION: Cyclic Analogs of Histatins
; FILE REFERENCE: 36555-0002

; CURRENT APPLICATION NUMBER: US/09/693,822B
; CURRENT FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: CA 2,285,673
; PRIOR FILING DATE: 1999-10-21
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: cyclic analogues of histatin
US-09-693-822B-2

Query Match 51.1%; Score 47; DB 4; Length 23;
Best Local Similarity 37.5%; Pred. No. 0.67;
Matches 6; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 SRHHCRSKAKRSRHH 16
||| : : :
Db 3 AKRHCHYKRFHEKHH 18
||| : : :||

RESULT 8
US-09-252-991A-28554
; Sequence 28554, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28554
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28554

Query Match 51.1%; Score 47; DB 4; Length 192;
Best Local Similarity 53.3%; Pred. No. 5.4;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 RRHHCRSKAKRSRHH 16
||| : : :||
Db 75 RQHCHTHAVRRHH 89
||| : : :||

RESULT 9
US-09-252-991A-23696
; Sequence 23696, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23696
; LENGTH: 618
; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23696

Query Match 51.1%; Score 47; DB 4; Length 618;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 RRHHCRSKAKRSRHH 15
||| : : :||
Db 166 RHHHLRARRRRQRH 179
||| : : :||

RESULT 10
US-09-252-991A-28358
; Sequence 28358, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28358
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28358

Query Match 51.1%; Score 47; DB 4; Length 618;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 RRHHCRSKAKRSRHH 15
||| : : :||
Db 166 RHHHLRARRRRQRH 179
||| : : :||

RESULT 11
US-09-693-822B-18
; Sequence 18, Application US/09693822B
; Patent No. 6555650
; GENERAL INFORMATION:
; APPLICANT: Lajoie, Gilles A.
; TITLE OF INVENTION: Cyclic Analogs of Histatins
; FILE REFERENCE: 36555-0002
; CURRENT APPLICATION NUMBER: US/09/693,822B
; CURRENT FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: CA 2,285,673
; PRIOR FILING DATE: 1999-10-21
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: cyclic analogues of histatin
US-09-693-822B-18

Query Match 50.0%; Score 46; DB 4; Length 19;
Best Local Similarity 42.9%; Pred. No. 0.78;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 RHHCRSKAKRSRHH 16
||| : : :||
Db 1 RHHCHYKRFHEKHH 14
||| : : :||

RESULT 14
US-09-724-864-45
; Sequence 45, Application US/09724864
; Patent No. 6380362
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Murison, James G.
; TITLE OF INVENTION: Polynucleotides, polypeptides expressed
; TITLE OF INVENTION: by the polynucleotides and methods for their use.
; FILE REFERENCE: 11000.1050U1

Search completed: May 21, 2004, 12:55:52
Job time : 24 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2004, 12:53:35 ; Search time 42 Seconds
(without alignments)
106.256 Million cell updates/sec

Title: US-09-933-780C-16

Perfect score: 92

Sequence: 1 SRRHCRSKAKRSRHH 16

Scoring table: BLOSUM62

Gapop 10.0 ; Gapext 0.5

Searched: 1149313 seqs, 278921704 residues

Total number of hits satisfying chosen parameters: 1149313

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US05_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92	100.0	16	11	US-09-933-780C-16
2	92	100.0	16	11	US-09-933-780C-39
3	92	100.0	16	15	US-10-353-678-3
4	89	96.7	16	11	US-09-933-780C-46
5	89	96.7	16	11	US-09-933-780C-50
6	88	95.7	16	11	US-09-933-780C-17
7	86	93.5	16	11	US-09-933-780C-40
8	86	93.5	16	11	US-09-933-780C-41
9	86	93.5	16	11	US-09-933-780C-45
10	86	93.5	16	11	US-09-933-780C-47
11	86	93.5	16	11	US-09-933-780C-48
12	86	93.5	16	11	US-09-933-780C-49
13	86	93.5	16	11	US-09-933-780C-51
14	83	90.2	16	11	US-09-933-780C-44
15	82	89.1	16	11	US-09-933-780C-42

16	82	89.1	16	11	US-09-933-780C-43	Sequence 43, Appl
17	72	78.3	13	11	US-09-933-780C-52	Sequence 52, Appl
18	60	65.2	14	15	US-10-353-678-4	Sequence 4, Appl
19	51.5	56.0	19723	15	US-10-084-846A-5	Sequence 5, Appl
20	49	53.3	511	12	US-10-267-502-349	Sequence 349, App
21	49	53.3	517	13	US-10-108-605-135	Sequence 135, App
22	49	53.3	543	12	US-10-425-114-64388	Sequence 64388, A
23	47	51.1	127	12	US-10-424-599-171717	Sequence 171717, A
24	47	51.1	150	12	US-10-425-114-63923	Sequence 63923, A
25	45.5	49.5	250	14	US-10-197-666A-78	Sequence 78, Appl
26	45.5	49.5	288	14	US-10-197-666A-80	Sequence 80, Appl
27	45	48.9	151	12	US-10-425-114-70414	Sequence 70414, A
28	45	48.9	255	10	US-09-866-050A-683	Sequence 683, App
29	45	48.9	484	14	US-10-278-481-13	Sequence 13, Appl
30	45	48.9	485	14	US-10-278-481-2	Sequence 2, Appl
31	44	47.8	63	12	US-10-391-068-23	Sequence 23, Appl
32	44	47.8	63	12	US-10-391-068-24	Sequence 24, Appl
33	44	47.8	63	12	US-10-391-068-25	Sequence 25, Appl
34	44	47.8	63	12	US-10-391-068-26	Sequence 26, Appl
35	44	47.8	179	12	US-10-424-599-216871	Sequence 216871, A
36	44	47.8	184	9	US-09-764-878-171	Sequence 171, App
37	44	47.8	184	14	US-10-079-854-171	Sequence 132, Ap
38	44	47.8	276	9	US-09-925-300-1392	Sequence 216870, A
39	44	47.8	291	12	US-10-424-599-216870	Sequence 2447, Ap
40	44	47.8	333	15	US-10-094-749-2447	Sequence 5769, Ap
41	44	47.8	431	14	US-10-106-698-5769	Sequence 20, Appl
42	44	47.8	435	15	US-10-362-445-20	Sequence 48783, A
43	43	46.7	99	9	US-09-864-761-48783	Sequence 59078, A
44	43	46.7	136	12	US-10-425-114-59078	Sequence 252610, A
45	43	46.7	166	12	US-10-424-599-252610	

ALIGNMENTS

RESULT 1

US-09-933-780C-16
; Sequence 16, Application US/09933780C
; Publication No. US20030229202A1
; GENERAL INFORMATION:
; APPLICANT: AVENTIS PHARMACEUTICALS INC.
; APPLICANT: GUO, Yong
; APPLICANT: MORSE, Clarence C
; APPLICANT: YAO, Zhengbin
; TITLE OF INVENTION: MEMBRANE PENETRATING PEPTIDES AND USES THEREOF
; FILE REFERENCE: HMR2053 PCT
; CURRENT APPLICATION NUMBER: US/09/933.780C
; CURRENT FILING DATE: 2001-08-21
; PRIOR FILING DATE: 2001-08-21
; PRIOR FILING DATE: 2000-08-25
; PRIOR FILING DATE: 2000-08-25
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 16
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Nuclear protein import sequence of hPER1
US-09-933-780C-16

Query Match 100.0%; Score 92; DB 11; Length 16;
Best Local Similarity 100.0%; Pred. No. 3.9e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRRHCRSKAKRSRHH 16
| | | | | | | | | | | | | | | |
DB 1 SRRHCRSKAKRSRHH 16

RESULT 2

US-09-933-780C-39

Current applications

```
; Sequence 39, Application US/09933780C
; Publication No. US20030229202A1
; GENERAL INFORMATION:
; APPLICANT: AVENTIS PHARMACEUTICALS INC.
; APPLICANT: GUO, Yong
; APPLICANT: MORSE, Clarence C
; APPLICANT: YAO, Zhengbin
; TITLE OF INVENTION: MEMBRANE PENETRATING PEPTIDES AND USES THEREOF
; FILE REFERENCE: HMR2053 PCT
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/227,647
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: GB 0103110.3
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 39
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
; US-09-933-780C-39

Query Match      100.0%; Score 92; DB 11; Length 16;
Best Local Similarity 100.0%; Pred. No. 3.9e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRRHCRSKAKRSRHH 16
   |||||:|||||
Db 1 SRRHCRSKAKRSRHH 16

RESULT 3
US-10-353-678-3
; Sequence 3, Application US/10353678
; Publication No. US20040002455A1
; GENERAL INFORMATION:
; APPLICANT: Uger, Bob
; APPLICANT: Salha, Danielle
; APPLICANT: Barber, Brian
; APPLICANT: Morse, Buzzy
; APPLICANT: Guo, Yong
; APPLICANT: Cheng, Su
; TITLE OF INVENTION: Targeted Immunogens
; FILE REFERENCE: API-01-018-US
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: US 10/353,678
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: US 60/352,892
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: US 10/219,850
; PRIOR FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 3
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Peptide (PERI-1) derived from protein transduction domain of
; US-10-353-678-3

Query Match      100.0%; Score 92; DB 15; Length 16;
Best Local Similarity 100.0%; Pred. No. 3.9e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRRHCRSKAKRSRHH 16
   |||||:|||||
Db 1 SRRHCRSKAKRSRHH 16
```

NFA

```
RESULT 4
US-09-933-780C-46
; Sequence 46, Application US/09933780C
; Publication No. US20030229202A1
; GENERAL INFORMATION:
; APPLICANT: AVENTIS PHARMACEUTICALS INC.
; APPLICANT: GUO, Yong
; APPLICANT: MORSE, Clarence C
; APPLICANT: YAO, Zhengbin
; TITLE OF INVENTION: MEMBRANE PENETRATING PEPTIDES AND USES THEREOF
; FILE REFERENCE: HMR2053 PCT
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/227,647
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: GB 0103110.3
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 46
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
; US-09-933-780C-46

Query Match      96.7%; Score 89; DB 11; Length 16;
Best Local Similarity 93.8%; Pred. No. 1e-05;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRRHCRSKAKRSRHH 16
   |||||:|||||
Db 1 SRRHCRSKAKRSRHH 16

RESULT 5
US-09-933-780C-50
; Sequence 50, Application US/09933780C
; Publication No. US20030229202A1
; GENERAL INFORMATION:
; APPLICANT: AVENTIS PHARMACEUTICALS INC.
; APPLICANT: GUO, Yong
; APPLICANT: MORSE, Clarence C
; APPLICANT: YAO, Zhengbin
; TITLE OF INVENTION: MEMBRANE PENETRATING PEPTIDES AND USES THEREOF
; FILE REFERENCE: HMR2053 PCT
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/227,647
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: GB 0103110.3
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 50
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
; US-09-933-780C-50

Query Match      96.7%; Score 89; DB 11; Length 16;
Best Local Similarity 93.8%; Pred. No. 1e-05;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRRHCRSKAKRSRHH 16
   |||||:|||||
Db 1 SRRHCRSKAKRSRHH 16

RESULT 6
US-09-933-780C-50
; Sequence 50, Application US/09933780C
; Publication No. US20030229202A1
; GENERAL INFORMATION:
; APPLICANT: AVENTIS PHARMACEUTICALS INC.
; APPLICANT: GUO, Yong
; APPLICANT: MORSE, Clarence C
; APPLICANT: YAO, Zhengbin
; TITLE OF INVENTION: MEMBRANE PENETRATING PEPTIDES AND USES THEREOF
; FILE REFERENCE: HMR2053 PCT
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/227,647
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: GB 0103110.3
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 50
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
; US-09-933-780C-50

Query Match      96.7%; Score 89; DB 11; Length 16;
Best Local Similarity 93.8%; Pred. No. 1e-05;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRRHCRSKAKRSRHH 16
   |||||:|||||
Db 1 SRRHCRSKAKRSRHH 16
```

```

: Sequence 41, Application US/09933780C
: Publication No. US20030229202A1
: GENERAL INFORMATION:
: APPLICANT: AVENTIS PHARMACEUTICALS INC.
: APPLICANT: GUO, Yong
: APPLICANT: MORSE, Clarence C
: APPLICANT: YAO, Zhengbin
: TITLE OF INVENTION: MEMBRANE PENETRATING PEPTIDES AND USES THEREOF
: FILE REFERENCE: HMR2053 PCT
: CURRENT APPLICATION NUMBER: US/09/933,780C
: CURRENT FILING DATE: 2001-08-21
: PRIOR APPLICATION NUMBER: US 60/227,647
: PRIOR FILING DATE: 2000-08-25
: PRIOR APPLICATION NUMBER: GB 0103110.3
: PRIOR FILING DATE: 2001-02-07
: NUMBER OF SEQ ID NOS: 54
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 41
: LENGTH: 16
: TYPE: PPT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Synthetic peptide
US-09-933-780C-41

```

Query Match 93.5%; Score 86; DB 11; Length 16;
Best Local Similarity 93.8%; Pred. No. 2.6e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 SRHHCRSKAKRSRHH 16
1 SRAHHCRSKAKRSRHH 16

RESULT 9
 US-09-933-780C-45
 ; Sequence 45, Application US/09933780C.
 ; Publication NO. US20030229202A1
 ; GENERAL INFORMATION:
 ; APPLICANT: AVENTIS PHARMACEUTICALS INC.
 ; APPLICANT: GUO Yong
 ; APPLICANT: MORSE, Clarence C
 ; APPLICANT: YAO, Zhengbin
 ; TITLE OF INVENTION: MEMBRANE PENETRATING PEPTIDES AND USES THEREOF
 ; FILE REFERENCE: HMR2053 PCT
 ; CURRENT APPLICATION NUMBER: US/09/933,780C
 ; CURRENT FILING DATE: 2001-08-21
 ; PRIOR APPLICATION NUMBER: US 60/227,647
 ; PRIOR FILING DATE: 2000-08-25
 ; PRIOR APPLICATION NUMBER: GB 0103110.3
 ; PRIOR FILING DATE: 2001-02-07
 ; NUMBER OF SEQ ID NOS: 54
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 45
 ; LENGTH: 16
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthetic peptide
 US-09-933-780C-45

Query Match 93.5%; Score 86; DB 11; Length 16;
Best Local Similarity 93.8%; Pred. No. 2.6e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 SRRHHCRSKAKRSRHH 16
|||||
1 SRRHHCASKAKRSRHH 16

RESULT 10
US-09-933-780C-47
; Sequence 47, Application US/099333780C

Publication No. US20030229202A1
; GENERAL INFORMATION:
; APPLICANT: AVENTIS PHARMACEUTICALS INC.
; APPLICANT: GUO, Yong
; APPLICANT: MORSE, Clarence C
; APPLICANT: YAO, Zhengbin
; TITLE OF INVENTION: MEMBRANE PENETRATING PEPTIDES AND USES THEREOF
; FILE REFERENCE: HMR2053 PCT
; CURRENT APPLICATION NUMBER: US/09/933,780C
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/227,647
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: GB 0103110.3
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 47
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-09-933-780C-47

Query Match 93.5%; Score 86; DB 11; Length 16;
Best Local Similarity 93.8%; Pred. No. 2.6e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SRRHCRSKAKRSRHH 16
|||||
DB 1 SRRHCRSAKRSRHH 16

RESULT 11
US-09-933-780C-48
; Sequence 48, Application US/09933780C
; Publication No. US20030229202A1
; GENERAL INFORMATION:
; APPLICANT: AVENTIS PHARMACEUTICALS INC.
; APPLICANT: GUO, Yong
; APPLICANT: MORSE, Clarence C
; APPLICANT: YAO, Zhengbin
; TITLE OF INVENTION: MEMBRANE PENETRATING PEPTIDES AND USES THEREOF
; FILE REFERENCE: HMR2053 PCT
; CURRENT APPLICATION NUMBER: US/09/933,780C
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/227,647
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: GB 0103110.3
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 48
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-09-933-780C-48

Query Match 93.5%; Score 86; DB 11; Length 16;
Best Local Similarity 93.8%; Pred. No. 2.6e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SRRHCRSKAKRSRHH 16
|||||
DB 1 SRRHCRSAKRSRHH 16

RESULT 12
US-09-933-780C-49
; Sequence 49, Application US/09933780C
; Publication No. US20030229202A1

GENERAL INFORMATION:
; APPLICANT: AVENTIS PHARMACEUTICALS INC.
; APPLICANT: GUO, Yong
; APPLICANT: MORSE, Clarence C
; APPLICANT: YAO, Zhengbin
; TITLE OF INVENTION: MEMBRANE PENETRATING PEPTIDES AND USES THEREOF
; FILE REFERENCE: HMR2053 PCT
; CURRENT APPLICATION NUMBER: US/09/933,780C
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/227,647
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: GB 0103110.3
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 49
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-09-933-780C-49

Query Match 93.5%; Score 86; DB 11; Length 16;
Best Local Similarity 93.8%; Pred. No. 2.6e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SRRHCRSKAKRSRHH 16
|||||
DB 1 SRRHCRSAKRSRHH 16

RESULT 13
US-09-933-780C-51
; Sequence 51, Application US/09933780C
; Publication No. US20030229202A1
; GENERAL INFORMATION:
; APPLICANT: AVENTIS PHARMACEUTICALS INC.
; APPLICANT: GUO, Yong
; APPLICANT: MORSE, Clarence C
; APPLICANT: YAO, Zhengbin
; TITLE OF INVENTION: MEMBRANE PENETRATING PEPTIDES AND USES THEREOF
; FILE REFERENCE: HMR2053 PCT
; CURRENT APPLICATION NUMBER: US/09/933,780C
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/227,647
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: GB 0103110.3
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-09-933-780C-51

Query Match 93.5%; Score 86; DB 11; Length 16;
Best Local Similarity 93.8%; Pred. No. 2.6e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SRRHCRSKAKRSRHH 16
|||||
DB 1 SRRHCRSAKRSRHH 16

RESULT 14
US-09-933-780C-44
; Sequence 44, Application US/09933780C
; Publication No. US20030229202A1
; GENERAL INFORMATION:

; APPLICANT: AVENTIS PHARMACEUTICALS INC.
; APPLICANT: GUO, Yong
; APPLICANT: MORSE, Clarence C
; APPLICANT: YAO, Zhengbin
; TITLE OF INVENTION: MEMBRANE PENETRATING PEPTIDES AND USES THEREOF
; FILE REFERENCE: HMR2053 PCT
; CURRENT APPLICATION NUMBER: US/09/933,780C
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/227,647
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: GB 0103110.3
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 44
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-09-933-780C-44

Query Match 90.2%; Score 83; DB 11; Length 16;
Best Local Similarity 93.8%; Pred. No. 6.7e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SRHHCRSKAKRSRHH 16
|||||
Db 1 SRHHARSKAKRSRHH 16

RESULT 15
US-09-933-780C-42
; Sequence 42, Application US/09933780C
; Publication No. US20030229202A1
; GENERAL INFORMATION:
; APPLICANT: AVENTIS PHARMACEUTICALS INC.
; APPLICANT: GUO, Yong
; APPLICANT: MORSE, Clarence C
; APPLICANT: YAO, Zhengbin
; TITLE OF INVENTION: MEMBRANE PENETRATING PEPTIDES AND USES THEREOF
; FILE REFERENCE: HMR2053 PCT
; CURRENT APPLICATION NUMBER: US/09/933,780C
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/227,647
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: GB 0103110.3
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 42
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-09-933-780C-42

Query Match 89.1%; Score 82; DB 11; Length 16;
Best Local Similarity 93.8%; Pred. No. 9.2e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SRHHCRSKAKRSRHH 16
|||||
Db 1 SRRAHCRSKAKRSRHH 16

Search completed: May 21, 2004, 12:56:47
Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2004, 12:49:34 ; Search time 11 Seconds
(without alignments)
75.738 Million cell updates/sec

Title: US-09-933-780C-16

Perfect score: 92

Sequence: 1 SRRHCRSKAKRSRHH 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	92	100.0	1290	1 PER1_HUMAN	O15534 homo sapien
2	88	95.7	1291	1 PER1_MOUSE	O35973 mus musculu
3	48	52.2	611	1 VIE3_MCMVS	P29832 murine cyto
4	45.5	49.5	288	1 TR2B_HUMAN	O15815 homo sapien
5	45	48.9	530	1 ESR2_RAT	O62986 rattus norv
6	44	47.8	61	1 HSP1_ANTLA	O8745 antechinomy
7	44	47.8	61	1 HSP1_SMIL0	Q9tuc2 smintchopsi
8	44	47.8	62	1 HSP1_MURLO	P42140 murexia lon
9	44	47.8	62	1 HSP1_SMIBI	Q9tuc4 smintchopsi
10	44	47.8	62	1 HSP1_SMIGR	Q9tuc3 smintchopsi
11	44	47.8	279	1 ESR2_MACMU	Q9tuc5 macaca mula
12	43	46.7	197	1 Y146_NPPOP	O10375 orygia pseu
13	43	46.7	247	1 LB38_ARATH	Q9en23 arabidopsis
14	43	46.7	530	1 ESR2_CALJA	Q95171 callithrix
15	43	46.7	530	1 ESR2_HUMAN	Q92731 homo sapien
16	43	46.7	530	1 ESR2_MOUSE	O08537 mus musculu
17	43	46.7	1425	1 MADI_HUMAN	O95405 homo sapien
18	42	45.7	102	1 HSP2_HYLLA	P35344 hylobates l
19	42	45.7	102	1 HSP2_MACMU	P35297 macaca mula
20	42	45.7	103	1 HSP2_MACNE	P34657 caenorhabdi
21	42	45.7	622	1 VP27_YEAST	P40343 saccharomyc
22	42	45.7	1198	1 MTR3_HUMAN	Q13615 homo sapien
23	41.5	45.1	219	1 TSF3_HELAN	P29675 helianthus
24	41.5	45.1	1038	1 SOG_DROME	Q24025 drosophila
25	41	44.6	63	1 HSP1_ANTST	P42129 antechinus
26	41	44.6	107	1 HSP2_MOUSE	P07978 mus musculu
27	41	44.6	266	1 YOTB_CABEL	P34657 caenorhabdi
28	41	44.6	481	1 CLK4_HUMAN	Q9haz1 homo sapien
29	40	43.5	102	1 HSP2_PONPY	P35301 pongo pygma
30	40	43.5	207	1 HCT2_CHLMU	P38020 chlamydia m
31	40	43.5	489	1 SFR4_MOUSE	Q8ve97 mus musculu
32	40	43.5	848	1 YBVB_YEAST	P38266 saccharomyc
33	40	43.5	851	1 YD72_SCHPO	Q10327 schizosacch

RESULT 1

ID	PER1_HUMAN	STANDARD;	PRT;	1290 AA.
AC	O15534;			
DT	15-JUL-1999 (Rel. 38, Created)			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Period circadian protein 1 (Circadian pacemaker protein Rigiui) (hPER).			
GN	PER1 OR PER OR RIGUI OR KIAA0482.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.			
RC	TISSUE=Heart;			
RX	MEDLINE=97462901; PubMed=9323128;			
RA	Sun Z.S., Albrecht U., Zhuchenko O., Bailey J., Eichele G., Lee C.C.;			
RT	"Rigiui, a putative mammalian ortholog of the Drosophila period gene.";			
RL	Cell 90:1003-1011(1997).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=97472418; PubMed=9333243;			
RA	Tei H., Okamura H., Shigeyoshi Y., Fukuhara C., Ozawa R., Hirose M.,			
RA	Sakaki Y.;			
RT	"Circadian oscillation of a mammalian homologue of the Drosophila			
RT	period gene.";			
RL	Nature 389:512-516(1997).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	Hida A., Sakaki Y., Tei H.;			
RT	"Genomic structures of the human and mouse periodl genes.";			
RL	Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Blood;			
RX	MEDLINE=20400339; PubMed=10940553;			
RA	Tarascio D., Zoraghi G.K., Falchi M., Iosi F., Paradisi S.,			
RA	Di Fiore B., Lavia P., Falbo V.;			
RT	"The human Perl gene: genomic organization and promoter analysis of			
RT	the first human orthologue of the Drosophila period gene.";			
RL	Gene 253:161-170(2000).			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RA	Nagase T., Kikuno R., Ohara O.;			
RT	"Homo sapiens cDNA clone from adult brain.";			
RL	Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.			
CC	-I- FUNCTION: Circadian regulator that may act as a transcription			
CC	factor. Behaves as a negative element in circadian transcriptional			
CC	loop. Does not appear to bind DNA, suggesting indirect			
CC	transcriptional inhibition. Seems to be the pacemaker component			
CC	which responds to light and mediates photic entrainment. In the			
CC	suprachiasmatic nucleus (SCN), it behaves like a day-type			
CC	oscillator, with maximum expression during the light period.			

34	40	43.5	1411	1	BEAL_HUMAN	Q15075 homo sapien
35	40	43.5	2052	1	FPV1_MOUSE	Q92156 mus musculu
36	39	42.4	363	1	GP78_HUMAN	Q96P69 homo sapien
37	39	42.4	448	1	WT1_RAT	P49952 rattus norv
38	39	42.4	449	1	WT1_MOUSE	P22561 mus musculu
39	39	42.4	484	1	SFRB_HUMAN	Q05519 homo sapien
40	39	42.4	605	1	VCLA_GOSHI	P09799 gossypium h
41	39	42.4	951	1	AHM3_ARATH	Q9szw4 arabidopsis
42	39	42.4	3718	1	LMAS_MOUSE	Q61001 mus musculu
43	38.5	41.8	67	1	YHFD_BACSU	O07602 bacillus su
44	38.5	41.8	247	1	MLF2_MOUSE	Q99KX1 mus musculu
45	38.5	41.8	248	1	MLF2_HUMAN	Q15773 homo sapien

ALIGNMENTS

Oscillations are maintained under constant darkness and are responsive to changes of the light/dark cycles. There is a 4 hour time delay between PER1 and PER2 oscillations. The expression rhythms appear to originate from retina (By similarity).

-1- SUBUNIT: Forms a heterodimer, probably with clock.

-1- SUBCELLULAR LOCATION: Nuclear (Potential).

-1- ALTERNATIVE PRODUCTS:

Event-Alternative splicing: Named isoforms=3;

Comment-Additional isoforms seem to exist;

Name=Rigui 4.7;

Isoid=O15534-1; Sequence=Displayed;

Name=Rigui 3.0;

Isoid=O15534-2; Sequence=Not described;

Name=Rigui 6.6; Synonyms=Truncated;

Isoid=O15534-3; Sequence=Not described;

-1- TISSUE SPECIFICITY: Widely expressed. Found in heart, brain, placenta, lung, liver, skeletal muscle, pancreas, and at low level in the kidney.

-1- SIMILARITY: Belongs to the basic helix-loop-helix (bHLH) family of transcription factors.

-1- SIMILARITY: Contains 2 PAS (PER-ARNT-SIM) dimerization domains.

-1- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

EMBL; AF022991; AAC51765.1; -

EMBL; AB002107; BAA22633.1; -

EMBL; AB030817; BAA94085.1; -

EMBL; AF102137; AAF15544.1; -

EMBL; AB088477; BAC06326.1; -

PIR; T00018; T00018.

Genew; HGNC:8845; PER1.

MIM; 602260; -

GO; GO:0009649; P:entrainment of circadian clock; TAS.

InterPro; IPR001610; PAS.

InterPro; IPR000014; PAS_domain.

Pfam; PF00989; PAS; 1.

SMART; SM00086; PAS; 1.

SMART; SM00091; PAS; 2.

PROSITE; PS0112; PAS; 1.

Transcription regulation; Nuclear protein; Repeat; Biological rhythms;

Alternative splicing.

DOMAIN 136 172 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).

DOMAIN 209 275 PAS 1.

DOMAIN 346 416 PAS 2.

DOMAIN 425 465 PAC.

DOMAIN 66 79 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).

DOMAIN 69 75 POLY-SER.

DOMAIN 653 656 POLY-SER.

DOMAIN 1010 1013 POLY-PRO.

DOMAIN 1269 1273 POLY-GLU.

DOMAIN 1276 1279 POLY-SER.

SEQUENCE 1290 AA; 136237 MW; 24BS3042869A4562 CRC64;

Query Match 100.0%; Score 92; DB 1; Length 1290;

Best Local Similarity 100.0%; Pred. No. 1.1e-06;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRHHCRSKAKSRHH 16

Db 830 SRHHCRSKAKSRHH 845

RESULT 2

PER1 MOUSE

ID PER1 MOUSE PRT; 1291 AA.

AC O35973;

DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Period circadian protein 1 (Circadian pacemaker protein Rigui) (mPER) (M-Rigui).

GN PER1 OR PER OR RIGUI.

DE Mus musculus (Mouse).

OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OC NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RX MEDLINE=97462901; PubMed=9333128;

RA Sun Z.S., Albrecht U., Zhuchenko O., Bailey J., Echele G., Lee C.C.; "Rigui, a putative mammalian ortholog of the Drosophila period gene.";

RL Cell 90:1003-1011(1997).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/c; TISSUE=Brain;

RX MEDLINE=97472418; PubMed=9333243;

RA Tei H., Okamura H., Shigeyoshi Y., Fukuhara C., Ozawa R., Hirose M., Sakaki Y.;

RT "Circadian oscillation of a mammalian homologue of the Drosophila period gene.";

RL Nature 389:512-516(1997).

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=20313889; PubMed=10857746;

RA Hida A., Koike N., Hirose M., Hattori M., Sakaki Y., Tei H.;

RT "The human and mouse period genes: five well-conserved E-boxes additively contribute to the enhancement of mper1 transcription.";

RL Genomics 65:224-233(2000).

CC -1- FUNCTION: Circadian regulator that may act as a transcription factor. Behaves as a negative element in circadian transcriptional loop. Does not appear to bind DNA, suggesting indirect transcriptional inhibition. Seems to be the pacemaker component which responds to light and mediates photic entrainment. In the suprachiasmatic nucleus (SCN), it behaves like a day-type oscillator, with maximum expression during the light period. Oscillations are maintained under constant darkness and are responsive to changes of the light/dark cycles. There is a 4 hour time delay between PER1 and PER2 oscillations. The expression rhythms appear to originate from retina.

CC -1- SUBUNIT: Forms a heterodimer, probably with clock.

CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).

CC -1- TISSUE SPECIFICITY: In brain, highest expression is observed in the SCN. Highly expressed in the pyramidal cell layer of the piriform cortex, the periventricular part of the caudate-putamen, many thalamic nuclei, and the granular layer of the cerebellar cortex. Weaker expression is detected in most area of the brain, including cortical and non cortical structures. Expression but no oscillations occurs in the glomerular and mitral cell layers of the olfactory bulb, the internal granular layer of the cerebellum, the cornu ammonis and dentate gyrus of the hippocampus, the cerebral and piriform cortices. Also found in heart, brain, spleen, lung, liver, skeletal muscle, testis, and at low level in the kidney.

CC -1- DEVELOPMENTAL STAGE: Expressed in the suprachiasmatic nucleus (SCN) during late fetal and early neonatal life.

CC -1- INDUCTION: By light exposure during subjective night, but not during subjective day. This induction might be the initial clock-specific molecular event for photic entrainment in the SCN.

CC -1- SIMILARITY: Belongs to the basic helix-loop-helix (bHLH) family of transcription factors.

CC -1- SIMILARITY: Contains 2 PAS (PER-ARNT-SIM) dimerization domains.

CC -1- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC	EMBL; L06816; AAA74505.1; -	NUCLEAR LOCALIZATION SIGNAL (POTENTIAL)
DR	EMBL; M77846; AAA45948.1; -	GLU-RICH (ACIDIC)
DR	FIR; A40835; EDESM.	GLU-RICH (ACIDIC)
DR	InterPro; IPR005028; Herpes_I2.3.	GLY/SER-RICH.
DR	Pfam; PF03361; Herpes_I2.3; 1.	ASP/GLU-RICH (ACIDIC)
KW	Transcription regulation; Activator; Early protein;	NUCLEAR LOCALIZATION SIGNAL (POTENTIAL)
KW	Phosphorylation; Alternative splicing	SER-RICH.
FT	DOMAIN 141 147	POLY-GLN.
FT	DOMAIN 166 177	
FT	DOMAIN 180 188	
FT	DOMAIN 193 278	
FT	DOMAIN 231 239	
FT	DOMAIN 279 285	
FT	DOMAIN 343 362	
FT	DOMAIN 366 370	
SQ	SEQUENCE 611 AA; 68103 MW; E7F10C8048536E14 CRC64;	

Query Match	52.2%	Score 48;	DB 1;	Length 611;
Best Local Similarity	56.2%;	Pred. No. 3.2;		
Matches 9;	Conservative 1;	Mismatches 6;	Indels 0;	Gaps 0;

SEQ	SEQUENCE	1291 AA; 136402 MW; A3DDCF706562937 CRC64;
1	SEQUENCE	1291 AA; 136402 MW; A3DDCF706562937 CRC64;

```

Query Match          95.7%; Score 88; DB 1: Length 1291;
Best Local Similarity 100.0%; Pred. No. 4.4e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 RHHHCCKAKRSRHH 16
        |||||
Db      828 RHHHCCKAKRSRHH 842

RESULT 3
VIE3_MCMVS
ID VIE3_MCMVS STANDARD; PRT; 611 AA.

```

(RA301),
SPRS10 OR TRA2B OR SILG41 OR SIG41.
Homo sapiens (Human),
Mus musculus (Mouse), and
Rattus norvegicus (Rat).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI TaxID=9606, 10090, 10116;
[1]_RN
SEQUENCES FROM N.A. (ISOFORMS 1 AND 2), AND SUBCELLULAR LOCATION.
SPECIES=Human; TISSUE=Cervical carcinoma;
MEDLINE=97355861; PubMed=9212162;
Beil B., Screaton G., Stamm S.;
"Molecular cloning of htra2-beta-1 and htra2-beta-2, two human
homologs of tra-2 generated by alternative splicing.";
DNA Cell Biol. 16:679-690(1997).
[2]_RN
SEQUENCE FROM N.A. (ISOFORM 3), AND TISSUE SPECIFICITY.
SPECIES=Human; RC
MEDLINE=99009334; PubMed=9790768;
Naylor O., Cap C., Stamm S.;
"Human transformer-2-beta gene (SPRS10): complete nucleotide sequence,
chromosomal localization, and generation of a tissue-specific
isoform.";
Genomics 53:191-202(1998).
[3]_RN
SEQUENCE FROM N.A. (ISOFORM 1).
SPECIES=Human; TISSUE=Brain;
Dauwalder B., Manzanares F.A., Mattox W.;
Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
[4]_RN
SEQUENCE FROM N.A. (ISOFORM 1).
RN

DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
DE Arginine/serine-rich splicing factor 10 (Transformer-2-beta) (HTRA2-
DE beta) (Transformer 2 protein homolog) (Silica-induced protein 41)
DE (RA301).
GN SFRS10 OR TRA2B OR SILG41 OR SIG41.
OS Homo sapiens (Human),
OS Mus musculus (Mouse), and
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxId=9606, 10090, 10116;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND SUBCELLULAR LOCATION.
RP SPECIES=Human; TISSUE=Cervical carcinoma;
RC MEDLINE=97355681; PubMed=9212162;
RX Beil B., Screaton G., Stamm S.;
EX "Molecular cloning of htra2-beta-1 and htra2-beta-2, two human
RT homologs of tra-2 generated by alternative splicing.";
RT DNA Cell Biol. 16:679-690(1997).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 3), AND TISSUE SPECIFICITY.
RP SPECIES=Human;
RC MEDLINE=99009334; PubMed=9790768;
RX Nayler C., Cap C., Stamm S.;
RT "Human transformer-2-beta gene (SFRS10): complete nucleotide sequence,
RT chromosomal localization, and generation of a tissue-specific
RT isoform.";
RT Genomics 53:191-202(1998).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RP SPECIES=Human; TISSUE=Brain;
RC Dauwalder B., Manzanares F.A., Mattox W.;
RX Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RP

SPCIRS=Human; TISSUE=Lung, Placenta, and Skin;
MEDLINE=2238257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Haddad S., Garcia A.M., Gay L.J., Rulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[5]
RN SEQUENCE FROM N.A., AND INDUCTION.
RP SPECIES=Mouse; TISSUE=Macrophage;
RX MEDLINE=95173444; PubMed=7868905;
RA Segade F., Claudio E., Wrobel K., Ramos S., Lazo P.S.;
RT "Isolation of nine gene sequences induced by silica in murine
macrophages.";
RL J. Immunol. 154:2384-2392 (1995).
[6]
RN SEQUENCE FROM N.A., AND INDUCTION.
RP SPECIES=Rat; STRAIN=Sprague-Dawley; TISSUE=Astocytes;
RX MEDLINE=96081858; PubMed=7499316;
RA Matsuo N., Ogawa S., Imai Y., Takagi T., Tohyama M., Stern D.,
RA Wanaka A.;
RT "Cloning of a novel RNA binding polypeptide (RA301) induced by
hypoxia/reoxygenation.";
RL J. Biol. Chem. 270:28216-28222 (1995).
[7]
RN FUNCTION, SUBCELLULAR LOCATION, AND PHOSPHORYLATION.
RP SPECIES=Human; TISSUE=Cervical carcinoma;
RX MEDLINE=98206475; PubMed=9546399;
RA Take R., Tohyama M., Ogawa S., Manley J.L.;
RT "Human Tra2 proteins are sequence-specific activators of pre-mRNA
splicing.";
RL Cell 53:139-148 (1998).
[8]
RN INTERACTION WITH SAFB/SAFB1.
RP SPECIES=Rat;
RX MEDLINE=98337913; PubMed=9671816;
RA Nayler O., Straetling W., Bourquin J.-P., Stagliar I., Lindemann L.,
RA Jasper H., Hartmann A.M., Fackelmeyer F.O., Ullrich A., Stamm S.;
RT "SAF-B couples transcription and pre-mRNA splicing to SAR/MAR
elements.";
RL Nucleic Acids Res. 26:3542-3549 (1998).
CC -!- FUNCTION: Sequence-specific RNA-binding protein which participates
in the control of pre-mRNA splicing.
CC -!- SUBUNIT: Binds to A3 enhancer proteins SRp75, SRp55, SRp40 and
SP30. Interacts with SAFB/SAFB1.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=3;
Name=1; Synonyms=HTRA2-beta1;
IsoId=Q15815-1; Sequence=Displayed;
Name=2; Synonyms=HTRA2-beta2;
IsoId=Q15815-2; Sequence=VSP_005897, VSP_005898, VSP_005899;
Note=Has been shown to exist only in human so far;
Name=3; Synonyms=HTRA2-beta3;
IsoId=Q15815-3; Sequence=VSP_005896;
Note=Has been shown to exist only in human so far;
CC -!- TISSUE SPECIFICITY: Highest expression in heart, skeletal muscle
and pancreas. Less abundant in kidney, placenta and brain. Lowest

expression in kidney and liver.
-!- INDUCTION: Induced by reoxygenation following hypoxia and by
exposure to silica. Repressed by interferon gamma, LPS and TPA.
-!- PTM: Phosphorylated in the RS domains.
-!- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
-!- SIMILARITY: Belongs to the SR family of splicing factors.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation in
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).

CC EMBL: U61267; AAC28242.1; -
CC EMBL: U68063; AAB08701.1; -
CC EMBL: U87836; AAB69763.1; -
CC EMBL: AF057159; AAD19277.1; -
CC EMBL: AF057159; AAD19278.1; -
CC EMBL: AF057159; AAD19279.1; -
CC EMBL: BC000160; AAH00160.1; -
CC EMBL: BC000451; AAH00451.1; -
CC EMBL: BC005898; AAH05898.1; -
CC EMBL: D49708; CAA56518.1; -
CC PIR: S68798; S68798.
CC HSPSP: P11940; ICDV.
CC Genew: HGNC:10781; SFRS10.
CC GK: Q15815; -
CC MIM: 602719; -
CC MGI: 106016; Silg41.
CC GO: GO:0005634; P:nucleus; IDA.
CC GO: GO:0008248; P:pre-mRNA splicing factor activity; IDA.
CC GO: GO:0000398; P:nuclear mRNA splicing, via spliceosome; IDA.
CC InterPro: IPR000504; RNA_rec_mot.
CC Pfam: PF000076; Rrm; 1.
CC SMART: SM00360; RRM; 1.
CC PROSITE: PS0102; RRM; 1.
CC PROSITE: PS00030; RRM_RNP_1; 1.
CC RNA-binding; mRNA splicing; Nuclear processing; Phosphorylation;
KW Alternative splicing; Nuclear protein.
FT DOMAIN 31 113 ARG/SER-RICH (RS1 DOMAIN).
FT DOMAIN 118 196 RNA-BINDING (RRM).
FT DOMAIN 193 230 LINKER.
FT DOMAIN 231 287 ARG/SER-RICH (RS2 DOMAIN).
FT VARSPPLIC 1 100 Missing (in isoform 3).
FT VARSPPLIC 13 20 ESRASRS -> VNVEEGKC (in isoform 2).
FT VARSPPLIC 23 38 /FTId=VSP_005897.
FT VARSPPLIC 39 288 AHGSGKGRHTPARSR -> RHLTSPFINEYKLRLNK (in
isoform 2).
FT VARSPPLIC Missing (in isoform 2).
FT VARSPPLIC /FTId=VSP_005899.
SQ SEQUENCE 288 AA; 33665 MW; 60B310C8BA443E28 CRC64;
Query Match 49.5%; Score 45.5; DB 1; Length 288;
Best Local Similarity 62.5%; Pred. No. 3.7;
Matches 10; Conservative 3; Mismatches 2; Indels 1; Gaps 1;
QY 1 SRRHHCRSKAKRSRSH 16
DB 64 SRRHYTRSRS-RSRSH 78
RESULT 5
ESR2 RAT
ID ESR2 RAT STANDARD; PRT; 530 AA.
AC Q62386; Q35784; Q35785; O55015; O55016; O70195; Q9R185;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Estrogen receptor beta (ER-beta).
GN ER2 OR NR3A2 OR ERBETA.
OS Rattus norvegicus (RAT).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
ON NCBI_TaxID=10116;
OX (1)
RP SEQUENCE OF 46-530 FROM N.A. (ISOFORM 1).
RA STRAIN=Sprague-Dawley; TISSUE=Prostate;
RX MEDLINE=96234066; PubMed=8650195;
RA Kuiper G.G.J.M., Enmark E., Peltö-Huikko M., Nilsson S.,
RA Gustafsson J.-A.;
RT "Cloning of a novel receptor expressed in rat prostate and ovary."
RT Proc. Natl. Acad. Sci. U.S.A. 93:5925-5930(1996).
RN (2)
RP SEQUENCE OF 46-530 FROM N.A. (ISOFORM 2).
RC STRAIN=Wistar; TISSUE=Ovary;
RX MEDLINE=98262932; PubMed=9600083;
RA Maruyama K., Endoh H., Sasaki-Iwaka H., Kanou H., Shimaya E.,
RA Hashimoto S., Kato S., Kawashima H.;
RT "A novel isoform of rat estrogen receptor beta with 18 amino acid
RT insertion in the ligand binding domain as a putative dominant
RT negative regulator of estrogen action."
RL Biochem. Biophys. Res. Commun. 246:142-147(1998).
RN (3)
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC STRAIN=Wistar; TISSUE=Prostate;
RA Aldridge T.C.;
RT "Tissue specific responses to estrogen: an explanation based on
RT differential activation of multiple estrogen receptors with different
RT estrogen response elements."
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN (4)
RP SEQUENCE OF 46-530 FROM N.A. (ISOFORMS 1; 3 AND 4).
RC STRAIN=Sprague-Dawley;
RX MEDLINE=98151005; PubMed=9492041;
RA Petersen D.N., Tkalec G.T., Koza-Taylor P.H., Turi T.G.,
RA Brown T.A.;
RT "Identification of estrogen receptor beta2, a functional variant of
RT estrogen receptor beta expressed in normal rat tissues."
RL Endocrinology 139:1082-1092(1998).
RN (5)
RP PARTIAL SEQUENCE FROM N.A. (ISOFORM 5).
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RA Price R., Handa R.J.;
RT "A novel splice variant of estrogen receptor beta found in rat
RT brain."
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Binds estrogens with an affinity similar to that of ER-
CC ALPHA, and activates expression of reporter genes containing
CC estrogen response elements (ERE) in an estrogen-dependent manner.
CC Isoform 3 and isoform 4 are unable to bind DNA and activate
CC transcription due to the truncation of the DNA binding domain.
CC Isoform 2 shows loss of ligand binding affinity and suppresses ER-
CC ALPHA and ER-BETA mediated transcriptional activation and may act
CC as a dominant negative regulator of estrogen action.
CC -!- SUBUNIT: Binds DNA as a homodimer. Can form a heterodimer with
CC ESR1. Can also form heterodimers between isoforms 1 and 2.
CC Interacts with NCOA3, NCOA5 and NCOA6 coactivators, leading to a
CC strong increase of transcription of target genes (by similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=5;
CC Comment=Additional isoforms seem to exist;
CC Name=1; Synonyms=Beta1;
CC IsoId=Q62986-1; Sequence=Displayed;
CC Name=2; Synonyms=Beta2;
CC IsoId=Q62986-2; Sequence=VSP_003699;
CC Name=3; Synonyms=Beta1-delta3;
CC IsoId=Q62986-3; Sequence=VSP_003697;
CC Name=4; Synonyms=Beta2-delta3;
CC IsoId=Q62986-4; Sequence=VSP_003697, VSP_003699;
CC Name=5; Synonyms=Beta1-delta4;
CC

CC IsoId=Q62986-5; Sequence=VSP_003698;
CC -!- TISSUE SPECIFICITY: EXPRESSED IN PROSTATE, OVARY, LIVER,
CC KIDNEY, FAT, BONE, BRAIN, UTERUS AND TESTIS.
CC -!- DOMAIN: Composed of three domains: a modulating N-terminal domain,
CC a DNA-binding domain and a C-terminal steroid-binding domain.
CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
CC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U57439; AAC52602.1; -;
CC EMBL; AB012721; BAA25431.1; -;
CC EMBL; AJ002602; CAA05631.1; ALT_INIT.
CC EMBL; AJ002603; CAA05632.1; -;
CC EMBL; AF042058; AAB97424.1; -;
CC EMBL; AF042059; AAB97425.1; -;
CC EMBL; AF042060; AAB97426.1; -;
CC EMBL; AF042061; AAB97427.1; -;
CC EMBL; AF161187; AAD47637.1; -;
CC PDB; 1HJ1; 04-JAN-02.
CC PDB; 1OKN; 28-JUL-00.
CC GO; GO:0005634; C:nucleus; ISS.
CC GO; GO:0030284; P:estrogen receptor activity; ISS.
CC GO; GO:0004879; P:ligand-dependent nuclear receptor activity; ISS.
CC GO; GO:0048019; P:receptor antagonist activity; ISS.
CC GO; GO:0005496; P:steroid binding; ISS.
CC GO; GO:0030520; P:estrogen receptor signaling pathway; ISS.
CC GO; GO:0030308; P:negative regulation of cell growth; ISS.
CC GO; GO:0006355; P:regulation of transcription, DNA-dependent; ISS.
CC InterPro; IPR001723; Stdrhm rec lig.
CC InterPro; IPR008946; Str ncl receptor.
CC InterPro; IPR001628; Znf_C4Steroid.
CC Pfam; PF00104; hormone_rec; 1.
CC Pfam; PF00105; zf-C4; 1.
CC PRINTS; PR00398; STRDHORMONER.
CC PRINTS; PR00047; STROIDFINGER.
CC ProDom; PD000035; Znf_C4steroid; 1.
CC SMART; SM00430; HOLI_1.
CC SMART; SM00399; Znf_C4; 1.
CC PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
CC Receptor; Transcription regulation; DNA-binding; Nuclear protein;
CC Zinc-finger; Steroid-binding; Alternative splicing; Phosphorylation;
CC 3D-structure.
CC
CC DOMAIN 1 148
CC DNA_BIND 149 214 NUCLEAR RECEPTOR-TYPE.
CC ZN_FING 149 169 C4-TYPE.
CC ZN_FING 185 209 C4-TYPE.
CC DOMAIN 215 530 STEROID-BINDING.
CC MOD_RES 87 87 PHOSPHORYLATION (BY SIMILARITY).
CC MOD_RES 488 488 PHOSPHORYLATION (BY SIMILARITY).
CC VARSPLIC 179 217 Missing (in isoform 3 and isoform 4).
CC VARSPLIC 219 318 Missing (in isoform 5).
CC VARSPLIC 364 364 /FTID=VSP_003698.
CC R -> RSSEDPHHVAQMKSAAPR (in isoform 2
CC and isoform 4).
CC /FTID=VSP_003699.
CC L -> Q (IN REF. 3 AND 4).
CC P -> A (IN REF. 3 AND 4).
CC S -> P (IN REF. 3; CAA05631).
CC S -> P (IN REF. 3; CAA05631).
CC SEQUENCE 530 AA; 59152 MW; 36F26D9DFD773DA9 CRC64;
CC

Query Match 48.9%; Score 45; DB 1; Length 530;
Best Local Similarity 56.2%; Pred. No. 8.1;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```

QY      1 SRHHCRSKAKSRHH 16
DB      236 SEQVHCLSKAKNGGH 251

RESULT 6
HSPI_ANTLA
ID HSPI_ANTLA STANDARD; PRT; 61 AA.
AC O18745;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Sperm protamine P1.
GN PRM1.
OS Antechinus laniger (Eastern jerboa marsupial).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Dasyuromorphia; Dasyuridae; Antechinomys.
OX NCBI_TaxID=60701;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97446380; PubMed=9299228;
RA Krajewski C., Blacket M., Buckley L., Westernman M.;
RT "A multigene assessment of phylogenetic relationships within the
RT dasyurid marsupial subfamily Sminthopsinae.";
RL Mol. Phylogenet. Evol. 8:236-248(1997).
CC -!- FUNCTION: Protamines substitute for histones in the chromatin of
CC sperm during the haploid phase of spermatogenesis. They compact
CC sperm DNA into a highly condensed, stable and inactive complex.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Testis.
CC -!- SIMILARITY: Belongs to the protamine P1 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF089881; AAD53340.1; -
CC InterPro; IPR000221; Protamine P1.
CC Pfam; PF00260; Protamine P1; 1.
CC PROSITE; PS00048; PROTAMINE_P1; 1.
CC Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
CC Testis; DNA condensation; Nuclear protein.
CC KW Testis; DNA condensation; Nuclear protein.
CC FT INIT MET 0 BY SIMILARITY.
CC SQ SEQUENCE 61 AA; 8409 MW; E021567627E562B1 CRC64;

Query Match 47.8%; Score 44; DB 1; Length 61;
Best Local Similarity 50.0%; Pred. No. 1.3;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      1 SRHHCRSKAKSRHH 16
DB      11 SRSRYRRRRRRSRHH 26

RESULT 8
HSPI_MURLO
ID HSPI_MURLO STANDARD; PRT; 62 AA.
AC P42140; P42150; P42154;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Sperm protamine P1.
GN PRM1.
OS Murexia longicauda (Short-furred dasyure),
OS Phascogale tapoatafa (Common wambenger),
OS Sminthopsis crassicauda (fat-tailed dunnart),
OS Myrmecobius fasciatus (Numbat), and
OS Thylacinus cynocephalus (Tasmanian wolf).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Dasyuromorphia; Dasyuridae; Murexia.
OX NCBI_TaxID=37736; 9293; 9301; 55782; 9275;
RN [1]
RP SEQUENCE FROM N.A.
RX SPECIES=M.longicauda, P.tapoatafa, and S.crassicauda;
RC TISSUE=Sperm;
RX MEDLINE=95215351; PubMed=7700877;
RA Retief J.D., Krajewski C., Westernman M., Winkfein R.J., Dixon G.H.;
RT "Molecular phylogeny and evolution of marsupial protamine P1 genes.";
RL Proc. R. Soc. Lond., B, Biol. Sci. 259:7-14(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX SPECIES=M.fasciatus, and T.cynocephalus;
RC MEDLINE=97368867; PubMed=9225481;
RA Krajewski C., Buckley L., Westernman M.;
RT "DNA phylogeny of the marsupial wolf resolved.";
RL Proc. R. Soc. Lond., B, Biol. Sci. 264:911-917(1997).
CC -!- FUNCTION: Protamines substitute for histones in the chromatin of

```

sperm during the haploid phase of spermatogenesis. They compact sperm DNA into a highly condensed, stable and inactive complex.

-1- SUBCELLULAR LOCATION: Nuclear.

-1- TISSUE SPECIFICITY: Testis.

-1- SIMILARITY: Belongs to the protamine P1 family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; L35336; AAA74600.1; -
EMBL; L35327; AAA74606.1; -
EMBL; L32743; AAA99478.1; -
EMBL; U87139; AAB91327.1; -
EMBL; U87140; AAB91328.1; -
InterPro; IPR000221; Protamine P1.
Pfam; PF00260; protamine P1; 1.
PROSITE; PS00048; PROTAMINE P1; 1.
Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding; Testis; DNA condensation; Nuclear protein.
INIT MET 0 BY SIMILARITY.
SEQUENCE 62 AA; 8566 MW; 99C02857CBB73429 CRC64;

Query Match 47.8%; Score 44; DB 1; Length 62;
Best Local Similarity 50.08; Pred. No. 1.3;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy I SRHHCSKAKRSRH 16
|| : : : :
Db 11 SRSRYRRRRRRSRH 26

RESULT 9

HSPI SMIBI ID HSP1 SMIBI STANDARD; PRT; 62 AA..

AC Q9TUC4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Sperm protamine P1.
GN PRM1.
OS Sminthopsis bindi (Kakadu dunnart).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Dasyuromorphia; Dasyuridae; Sminthopsis.
OC NCBI_taxid=90757;
RN [1]
RP SEQUENCE FROM N.A.
RX BLACKET=99310778; PubMed=10381317;
RA Medline M.J., Krajewski C., Labrinidis A., Cambron B., Cooper S.,
RA Westerman M.;
RT Systematic relationships within the dasyurid marsupial tribe
RT Sminthopini -- a multigene approach.;
RL Mol. Phylogenet. Evol. 12:140-155 (1999).
CC -1- FUNCTION: Protamines substitute for histones in the chromatin of
CC sperm during the haploid phase of spermatogenesis. They compact
CC sperm DNA into a highly condensed, stable and inactive complex.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: Testis.
CC -1- SIMILARITY: Belongs to the protamine P1 family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF089873; AAD55332.1; -

```

ID ESR2 MACMU STANDARD; PRT; 279 AA.
AC Q9TTE5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-SEP-2003 (Rel. 41, Last annotation update)
DE Estrogen receptor beta (ER-beta) (Fragment).
GN ESR2 OR NR3A2.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20113291; PubMed=10644527;
RA Wu X.X., Ma X.H., Smith G.C.S., Nathaniellez P.W.;
RT "Differential distribution of Eralpha and ERbeta mRNA in intrauterine
RL tissues of the pregnant rhesus monkey."
RL Am. J. Physiol. 278:C190-C198(2000).
CC -!- FUNCTION: Nuclear hormone receptor. Binds estrogens with an
CC affinity similar to that of ESR1 (ER-alpha), and activates
CC expression of reporter genes containing estrogen response elements
CC (ERE) in an estrogen-dependent manner. May play a role in ovarian
CC follicular growth and maturation.
CC -!- SUBUNIT: Binds DNA as a homodimer. Can form a heterodimer with
CC ESR1. Interacts with NCOA3, NCOA5 and NCOA6 coactivators, leading
CC to a strong increase of transcription of target genes (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- DOMAIN: Composed of three domains: a modulating N-terminal domain,
CC a DNA-binding domain and a C-terminal steroid-binding domain.
CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
CC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF119229; AAD54069.1; -.
CC HSSP; P03372; IERR.
DR InterPro; IPR000536; Hormone_rec_lig.
DR InterPro; IPR001723; Steroid_rcptor.
DR InterPro; IPR008946; Str_ncl_rcptor.
DR InterPro; IPR001628; Znfc4steroid.
DR Pfam; PF00104; hormone_rc; 1.
DR PRINTS; PR00398; STRDHORMNER.
DR SMART; SM00430; HOLLI; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; PARTIAL.
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW Zinc-finger; Steroid-binding; Phosphorylation.
FT NON_TER 1 1 STEROID-BINDING.
FT MOD_RES 251 251 PHOSPHORYLATION (BY SIMILARITY).
FT NON_TER 279 279
SQ SEQUENCE 279 AA; 31105 MW; 858D9B7D01DA0301 CRC64;

Query Match 47.8%; Score 44; DB 1; Length 279;
Best Local Similarity 66.7%; Pred. No. 6.1;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 HCRSKAKRSRHH 16
DB 3 HCAGKAKRSRSH 14

RESULT 12
Y146_NPVOP
ID Y146_NPVOP STANDARD; PRT; 197 AA.
AC O10375;

DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 21.8 kDa protein (ORF144).
OS Orygia pseudotsugata multicapsid polyhedrosis virus (OpNPV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=164623;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97271300; PubMed=9126251;
RA Ahrens C.H., Russell R.R., Funk C.J., Evans J., Harwood S.,
RA Rohrmann G.F.;
RT "The sequence of the Orygia pseudotsugata multinucleocapsid nuclear
RL polyhedrosis virus genome."
RL Virology 229:381-393(1997).
CC -!- SIMILARITY: TO CORRESPONDING ORF IN ACNPV.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U75930; AAC59143.1; -.
CC Hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 197 AA; 21797 MW; D5E4B5B585F79EBE CRC64;

Query Match 46.7%; Score 43; DB 1; Length 197;
Best Local Similarity 72.7%; Pred. No. 6.1;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 HCRSKAKRSRHH 15
DB 113 HRRSEAKRTRH 123

RESULT 13
LB38_ARATH
ID LB38_ARATH STANDARD; PRT; 247 AA.
AC Q9SN23; Q8LDW4;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE LOB domain protein 38.
GN LBD38 OR AT3G49940 OR F3A4.20.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosida II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016720; PubMed=11130713;
RA Salanoubat M., Lemcke K., Rieger M., Ansoorge W., Unseld M.,
RA Farman B., Valle G., Bloeker H., Perez-Alonso M., Obermaier B.,
RA Deleney M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,
RA De Simone V., Choinsne N., Artiguenave F., Robert C., Brottier P.,
RA Wincker P., Cattolico L., Weissenbach J., Saurin W., Quetier F.,
RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
RA Wurnbach E., Drzonek H., Erfle H., Jordan N., Bangert S.,
RA Wiedemann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
RA Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simonati B.,
RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordiek G.,
RA Reichelt J., Scharfe M., Schoen O., Barques M., Terol J., Climent J.,
RA Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,
RA Cooke R., Laudie M., Berger-Llauro C., Fumelle B., Masuy D.,
RA de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
RA Monfort A., Argiriou A., Flores M., Liguori R., Vitale D.,
RA Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
```

RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA Rooney T., Rizzo M., Walts A., Utterback T., Fujii C.Y., Shea T.P.,
RA Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
RA Pai G., Militscher J., Sellers P., Gill J.E., Feldblyum T.V.,
RA Preuss D., Lin X., Niernan W.C., Salzberg S.L., White O., Venter J.C.,
RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato I., Asamizu E.,
RA Sasamoto S., Kimura T., Idegawa K., Kawashima K., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsumoto A., Muraki A.,
RA Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Tabata S.,
RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:820-822(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=22954850; PubMed=14593172;
RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,
RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
RA Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,
RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
RA Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Empirical analysis of transcriptional activity in the Arabidopsis
RT genome.";
RL Science 302:842-846(2003).
RN [3]
RP SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.A.;
RT "Full-length cDNA from Arabidopsis thaliana.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP TISSUE SPECIFICITY
RX MEDLINE=22063719; PubMed=12068116;
RA Shuai B., Reynaga-Pena C.G., Springer P.S.;
RT "The LATERAL ORGAN BOUNDARIES gene defines a novel, plant-specific
RT gene family.";
RL Plant Physiol. 129:747-761(2002).
CC -!- TISSUE SPECIFICITY: Expressed in young shoots, roots, stems,
CC leaves and flowers.
CC -!- SIMILARITY: Contains 1 LOB domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL132378; CAB62102.1; -
CC DR EMBL; BT0002449; AAC00809.1; -
CC DR EMBL; AY085761; AAM62379.1; -
CC DR PIR; T45847; T45847.
CC DR InterPro; IPR004883; DUF260.
CC DR Pfam; PF03195; DUF260; 1.
CC DR PROSITE; PSS0891; LOB; 1.
CC DR DOMAIN 1 107 LOB.
CC FT DOMAIN 232 237 POLY-GLY.
CC FT CONFLICT 149 149 R -> Q (IN REF. 3).
CC FT SEQUENCE 247 AA; 26722 MW; 29BD9024481C1788 CRC64;
Query Match 46.7%; Score 43; DB 1; Length 247;
Best Local Similarity 63.6%; Pred. No. 7.6;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 HHCRRKAKRSR 14
Db 153 HHCRRSSRSR 163

RESULT 14
ESR2_CALJA STANDARD; PRT; 530 AA.
ID ESR2_CALJA
AC Q95171;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Estrogen receptor beta (ER-beta).
GN ESR2 OR NR3A2.
OS Callitrix jacchus (Common marmoset).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;
OC Callitrich.
OX NCBI_TaxID=9483;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Gaughan J., Scobie G.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Nuclear hormone receptor. Binds estrogens with an
CC affinity similar to that of ESR1 (ER-alpha), and activates
CC expression of reporter genes containing estrogen response elements
CC (ERE) in an estrogen-dependent manner. May play a role in ovarian
CC follicular growth and maturation.
CC -!- SUBUNIT: Binds DNA as a homodimer. Can form a heterodimer with
CC ESR1. Interacts with NCOA3, NCOA5 and NCOA6 coactivators, leading
CC to a strong increase of transcription of target genes (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- DOMAIN: Composed of three domains: a modulating N-terminal domain,
CC a DNA-binding domain and a C-terminal steroid-binding domain.
CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Y09372; CAA70546.2; -
CC DR HSSP; P03372; 1HCQ.
CC DR GO; GO:0005634; C:nucleus; ISS.
CC DR GO; GO:0030284; P:estrogen receptor activity; ISS.
CC DR GO; GO:0004879; F:ligand-dependent nuclear receptor activity; ISS.
CC DR GO; GO:0048019; F:receptor antagonist activity; ISS.
CC DR GO; GO:0005496; F:steroid binding; ISS.
CC DR GO; GO:0030520; P:estrogen receptor signaling pathway; ISS.
CC DR GO; GO:0030308; P:negative regulation of cell growth; ISS.
CC DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; ISS.
CC DR InterPro; IPR005336; Hormone_rec_lig.
CC DR InterPro; IPR001723; Steroid_receptor.
CC DR InterPro; IPR008946; Str_ncl_receptor.
CC DR InterPro; IPR001628; Znf_C4steroid.
CC DR Pfam; PF00104; hormone_rec; 1.
CC DR PRINTS; PR00398; STRDHOMNER.
CC DR PRINTS; PR00047; STROIDFINGER.
CC DR ProDom; PD000035; Znf_C4steroid; 1.
CC DR SMART; SM00430; HOL1; 1.
CC DR SMART; SM00399; Znf_C4; 1.
CC DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW Receptor; transcription regulation; DNA-binding; Nuclear protein;
KW Zinc-finger; Steroid-binding; Phosphorylation.
FT DOMAIN 1 148 MODULATING.
FT DNA_BIND 149 214 NUCLEAR RECEPTOR-TYPE.

FT ZN FLNG 149 169 C4-TYPE.
 FT ZN FLNG 185 209 C4-TYPE.
 FT DOMA14 215 530 STEROID-BINDING.
 FT MOD RES 87 87 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD RES 488 488 PHOSPHORYLATION (BY SIMILARITY).
 SQ SEQUENCE 530 AA; 59087 MW; 45D99107A84C53D1 CRC64;
 Query Match 46.7%; Score 43; DB 1; Length 530;
 Best Local Similarity 66.7%; Pred. No. 16;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 5 HCBKAKRGRHH 16
 DB 240 HCAGKAKRSGGH 251
 RESULT 15
 ESR2_HUMAN STANDARD; PRT; 530 AA.
 ID ESR2_HUMAN AC Q92731; O60608; O60685; O60702; O60703; O75583; O75584; Q9UEV6;
 AC Q9UHD3; Q9UOK9;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Estrogen receptor beta (ER-beta).
 GN ESR2 OR NR3A2 OR ESRB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=98139878; PubMed=9473491;
 RA Ogawa S., Inoue S., Watanabe T., Hiroi H., Orito A., Hosoi T.,
 RA Ouchi Y., Muramatsu M.;
 RT "The complete primary structure of human estrogen receptor beta
 RT (hERbeta) and its heterodimerization with ER alpha in vivo and in
 RT vitro.";
 RL Biochem. Biophys. Res. Commun. 243:122-126(1998).
 RN [2]
 RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
 RC TISSUE=Breast, and Testis;
 RX MEDLINE=98300286; PubMed=9636657;
 RA Moore J.T., McKee D.D., Slentz-Kesler K., Moore L.B., Jones S.A.,
 RA Horne E.L., Su J.-L., Kiewer S.A., Lehmann J.M., Willson T.M.;
 RT "Cloning and characterization of human estrogen receptor beta
 RT isoforms.";
 RL Biochem. Biophys. Res. Commun. 247:75-78(1998).
 RN [3]
 RP SEQUENCE OF 48-530 FROM N.A. (ISOFORM 1), AND CHARACTERIZATION.
 RC TISSUE=Testis;
 RX MEDLINE=96354875; PubMed=8769313;
 RA Moseelman S., Folman J., Dijkema R.;
 RT "ER beta: identification and characterization of a novel human
 RT estrogen receptor.";
 RL FEBS Lett. 392:49-53(1996).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 2), AND CHARACTERIZATION.
 RC TISSUE=Testis;
 RX MEDLINE=98337908; PubMed=9671811;
 RA Ogawa S., Inoue S., Watanabe T., Orito A., Hosoi T., Ouchi Y.,
 RA Muramatsu M.;
 RT "Molecular cloning and characterization of human estrogen receptor
 RT beta cx: a potential inhibitor of estrogen action in human.";
 RL Nucleic Acids Res. 26:3505-3512(1998).
 RN [5]
 RP SEQUENCE OF 59-530 FROM N.A. (ISOFORMS 7 AND 8).
 RC TISSUE=Endometrium;
 RA Brandenberger A.W., Lebovic D., Taylor R.N., Jaffe R.B.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM 3).
 RC TISSUE=Ovary;

RX MEDLINE=98348389; PubMed=9685228;
 RA Lu B., Leygue E., Dotzlaw H., Murphy L.J., Murphy L.C., Watson P.H.;
 RT "Estrogen receptor-beta mRNA variants in human and murine tissues.";
 RL Mol. Cell. Endocrinol. 138:199-203(1998).
 RN [7]
 RP SEQUENCE OF 1-69 FROM N.A.
 RA Li L.C., Dahlia R.;
 RT "Cloning and characterization of the estrogen receptor beta gene
 RT promoter.";
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP CHARACTERIZATION.
 RX MEDLINE=97467383; PubMed=9325313;
 RA Pace P., Taylor J., Suntharalingam S., Coombes R.C., Ali S.;
 RT "Human estrogen receptor beta binds DNA in a manner similar to and
 RT dimerizes with estrogen receptor alpha.";
 RL J. Biol. Chem. 272:25832-25838(1997).
 RN [9]
 RP INTERACTION WITH NCOA3.
 RX MEDLINE=97410321; PubMed=9267036;
 RA Chen H., Lin R.J., Schiltz R.L., Chakravarti D., Nash A., Nagy L.,
 RA Privaleky M.L., Nakatani Y., Evans R.M.;
 RT "Nuclear receptor coactivator ACTR is a novel histone
 RT acetyltransferase and forms a multimeric activation complex with p/CAF
 RT and CBP/p300.";
 RL Cell 90:569-580(1997).
 RN [10]
 RP INTERACTION WITH NCOA6.
 RX MEDLINE=20148724; PubMed=10681503;
 RA Cairra F., Antonson P., Pelto-Huikko M., Treuter E., Gustafsson J.-A.;
 RT "Cloning and characterization of RAP250, a nuclear receptor
 RT coactivator.";
 RL J. Biol. Chem. 275:5308-5317(2000).
 RN [11]
 RP INTERACTION WITH NCOA5.
 RX MEDLINE=20565767; PubMed=11113208;
 RA Sauve F., McBroome L.D.B., Gallant J., Moraitis A.N., Labrie F.,
 RA Giguere V.;
 RT "CIA, a novel estrogen receptor coactivator with a bifunctional
 RT nuclear receptor interacting determinant.";
 RL Mol. Cell. Biol. 21:343-353(2001).
 CC -1- FUNCTION: Nuclear hormone receptor. Binds estrogens with an
 CC affinity similar to that of ESR1, and activates expression of
 CC reporter genes containing estrogen response elements (ERE) in an
 CC estrogen-dependent manner. Isoform beta-cx lacks ligand binding
 CC ability and has no or only very low ere binding activity resulting
 CC in the loss of ligand-dependent transactivation ability. DNA-
 CC binding by ESR1 and ESR2 is rapidly lost at 37 degrees Celsius in
 CC the absence of ligand while in the presence of 17 beta-estradiol
 CC and 4-hydroxy-tamoxifen loss in DNA-binding at elevated
 CC temperature is more gradual.
 CC -1- SUBUNIT: Binds DNA as a homodimer. Can form a heterodimer with
 CC ESR1. Isoform beta-2/cx preferentially forms a heterodimer with
 CC ESR1 rather than ESR2 and inhibits DNA-binding by ESR1. Interacts
 CC with NCOA3, NCOA5 and NCOA6 coactivators, leading to a strong
 CC increase of transcription of target genes.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=8;
 CC Name=1; Synonyms=Beta-1;
 CC IsoId=Q92731-1; Sequence=Displayed;
 CC Name=2; Synonyms=Beta-2, CX;
 CC IsoId=Q92731-2; Sequence=VSP_003689;
 CC Name=3; Synonyms=Beta-2A;
 CC IsoId=Q92731-3; Sequence=VSP_003684, VSP_003686;
 CC Name=4; Synonyms=Beta-3;
 CC IsoId=Q92731-4; Sequence=VSP_003690;
 CC Name=5; Synonyms=Beta-4;
 CC IsoId=Q92731-5; Sequence=VSP_003691;
 CC Name=6; Synonyms=Beta-5;
 CC IsoId=Q92731-6; Sequence=VSP_003692;
 CC Name=7; Synonyms=Beta-5A;
 CC IsoId=Q92731-7; Sequence=VSP_003685;

CC Name=8; Synonyms=Beta-6;
CC IsoId=Q92731-8; Sequence=VSP_003687, VSP_003688;
CC -!- TISSUE SPECIFICITY: Isoform beta-1 is expressed in testis and
CC ovary, and at a lower level in heart, brain, placenta, liver,
CC skeletal muscle, spleen, thymus, prostate, colon, bone marrow, and
CC mammary gland and uterus. Also found in uterine bone, breast, and
CC ovarian tumor cell lines, but not in colon and liver tumors.
CC Isoform beta-2 is expressed in spleen, thymus, testis and ovary
CC and at a lower level in skeletal muscle, prostate, colon, small
CC intestine, leukocytes, bone marrow, mammary gland and uterus.
CC Isoform beta-3 is found in testis. Isoform beta-4 is expressed in
CC testis, and at a lower level in spleen, thymus, ovary, mammary
CC gland and uterus. Isoform beta-5 is expressed in testis, placenta,
CC skeletal muscle, spleen and leukocytes, and at a lower level in
CC heart, lung, liver, kidney, pancreas, thymus, prostate, colon,
CC small intestine, bone marrow, mammary gland and uterus. Not
CC expressed in brain.
CC -!- DOMAIN: Composed of three domains: a modulating N-terminal domain,
CC a DNA-binding domain and a C-terminal steroid-binding domain.
CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB006590; BAA24953.1; -
CC EMBL; AF051427; AAC05985.1; -
CC EMBL; AF051428; AAC05751.1; -
CC EMBL; AF061054; AAC39784.1; -
CC EMBL; AF061055; AAC39785.1; -
CC EMBL; AF060555; AAC15234.1; -
CC EMBL; X99101; CAA67555.1; ALT_INIT.
CC EMBL; AB006589; BAA31966.1; -
CC EMBL; AF074598; AAC25602.1; -
CC EMBL; AF074599; AAC25603.1; -
CC EMBL; AF124790; AAD32580.1; -
CC EMBL; AF047463; AAC03786.1; -
CC EMBL; AF191544; AAF24232.1; -
CC PDB; 1L2J; 01-MAY-02.
CC PDB; INDE; 18-DEC-02.
CC PDB; 1OKM; 28-JUL-00.
CC TRANSFAC; T04651; -
CC TRANSFAC; T05387; -
CC TRANSFAC; T05388; -
CC TRANSFAC; T05389; -
CC TRANSFAC; T05390; -
CC TRANSFAC; T05391; -
CC Genew; HGNC:3468; ESR2.
CC MIM; 601663; -
CC GO; GO:0005634; C:nucleus; TAS.
CC GO; GO:0030284; F:estrogen receptor activity; TAS.
CC GO; GO:0048019; F:receptor antagonist activity; NAS.
CC GO; GO:0005496; F:steroid binding; TAS.
CC GO; GO:0003713; F:transcription co-activator activity; TAS.
CC GO; GO:0003700; F:transcription factor activity; TAS.
CC GO; GO:0007267; P:cell-cell signaling; TAS.
CC GO; GO:0030520; P:estrogen receptor signaling pathway; TAS.
CC GO; GO:0030308; P:negative regulation of cell growth; NAS.
CC GO; GO:0006355; P:regulation of transcription, DNA-dependent; TAS.
CC GO; GO:0007165; P:signal transduction; TAS.
CC InterPro; IPR000536; Hormone_rec_lig.
CC InterPro; IPR001723; Steroid_receptor.
CC InterPro; IPR008946; Str_ncl_receptor.
CC InterPro; IPR001628; Znf_C4steroid.
CC Pfam; PF00104; hormone_rec; 1.
CC Pfam; PF00105; zf-C4; 1.
CC PRINTS; PR00398; STRDHORMONER.

Query Match 46.7%; Score 43; DB 1; Length 530;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 5 HCRSKAKGRSRHH 16
Db 240 HCAGKAKGRSGGH 251
Search completed: May 21, 2004, 12:53:51
Job time : 12 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2004, 12:50:15 ; Search time 39 Seconds
(without alignments)
129.443 Million cell updates/sec

Title: US-09-933-780C-16
Perfect score: 92
Sequence: 1 SRHHCKSKAKRSRH 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88	95.7	1244	11 Q8CHI5	Q8chi5 rattus norv
2	88	95.7	1271	11 Q8CHR1	Q8chr1 mus musculu
3	88	95.7	1285	11 Q8X3T3	Q8x3t3 spalax juda
4	51.5	56.0	341	5 Q8SX50	Q8sx50 drosophila
5	51.5	56.0	375	5 Q9W039	Q9w039 drosophila
6	50	54.3	179	16 Q8G4B1	Q8g4b1 bifidobacte
7	49	53.3	162	11 Q9D3A7	Q9d3a7 mus musculu
8	48	52.2	499	6 Q40367	Q40367 medicago sa
9	48	52.2	499	6 Q95WF0	Q95wf0 macaca arc
10	47	51.1	129	10 Q8LM94	Q8lm94 oryza sativ
11	47	51.1	340	5 Q95Q04	Q95q04 caenorhabdi
12	47	51.1	375	17 Q30258	Q30258 archaeoglob
13	47	51.1	840	10 Q9LN46	Q9ln46 arabidopsis
14	46.5	50.5	347	13 Q7SXU9	Q7sxu9 brachydanio
15	46	50.0	135	16 Q7UUR6	Q7uur6 rhodopirell
16	46	50.0	142	10 Q8GVX8	Q8gvx8 oryza sativ

17	45.5	49.5	146	6 Q8WMQ0	Q8wmq0 macaca mula
18	45.5	49.5	289	13 Q9DDU8	Q9ddu8 gallus gall
19	45	48.9	135	10 Q9CAH6	Q9cah6 arabidopsis
20	45	48.9	236	10 Q9LM00	Q9lm00 pinus taeda
21	45	48.9	282	10 Q8LQCS	Q8lqcs oryza sativ
22	45	48.9	342	5 Q8IKU9	Q8iku9 plasmodium
23	45	48.9	422	11 Q9D8T1	Q9d8t1 mus musculu
24	45	48.9	543	5 Q9VES0	Q9ves0 drosophila
25	44.5	48.4	58	16 Q8VJ97	Q8vj97 mycobacteri
26	44.5	48.4	327	5 Q95Y35	Q95y35 caenorhabdi
27	44.5	48.4	874	5 Q9VGG3	Q9vgg3 drosophila
28	44.5	48.4	893	5 Q8B999	Q8b999 drosophila
29	44	47.8	334	2 Q93L31	Q93l31 burkholderi
30	44	47.8	364	16 Q8UF36	Q8uf36 agrobacteri
31	44	47.8	366	10 Q43611	Q43611 phoenix dac
32	44	47.8	1165	5 Q86HN1	Q86hn1 dictyosteli
33	43.5	47.3	238	12 Q91TUS	Q91tus tupaia herp
34	43.5	47.3	278	13 Q7ZUG9	Q7zug9 brachydanio
35	43.5	47.3	306	13 Q7ZTW8	Q7ztw8 xenopus lae
36	43.5	47.3	438	5 Q9W3X8	Q9w3x8 drosophila
37	43.5	47.3	580	5 Q8IQH1	Q8iqh1 drosophila
38	43.5	47.3	635	5 Q96W9	Q96w9 drosophila
39	43.5	47.3	889	10 Q9FHH9	Q9fhh9 arabidopsis
40	43.5	47.3	1049	10 Q84UQ1	Q84uq1 oryza sativ
41	43	46.7	145	5 Q23606	Q23606 caenorhabdi
42	43	46.7	178	11 Q8CG84	Q8cg84 mus musculu
43	43	46.7	290	5 Q8TOF4	Q8tof4 drosophila
44	43	46.7	309	5 Q9VL63	Q9vl63 drosophila
45	43	46.7	323	4 Q86Z31	Q86z31 homo sapien

ALIGNMENTS

RESULT 1

Q8CHI5	PRELIMINARY;	PRT; 1244 AA.
ID Q8CHI5		
AC Q8CHI5		
DT 01-MAR-2003 (TrEMBLrel. 23, Created)		
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)		
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
DE Period1 (Fragment).		
GN RPER1.		
OS Rattus norvegicus (Rat).		
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX NCBI_TaxID=10116;		
RN [1]		
RP SEQUENCE FROM N.A.		
RC TISSUE=Brain;		
RA Suzuki S., Oishi K., Sakamoto K., Ishida N.;		
RT "Cloning and circadian expression of rat period3 gene.;"		
RL Submitted (Oct-2002) to the EMBL/GenBank/DBJ databases.		
DR EMBL; AB029276; BAC53666.1; -		
DR GO; GO:0004871; F:signal transduction; IEA.		
DR GO; GO:0007165; P:signal transducer activity; IEA.		
DR InterPro; IPR000014; PAS domain.		
DR InterPro; IPR002965; P-rich_extensn.		
DR Pfam; PF00989; PAS; 1.		
DR PRINTS; PR01217; PRICHEXTENSN.		
DR SMART; SM00091; PAS; 2.		
DR PROSITE; PS01112; PAS; 1.		
FT NON_TER 1244 1244		
SQ SEQUENCE 1244 AA; 131670 MW; 11D67FC14E88062D CRC64;		

Query Match 95.7%; Score 88; DB 11; Length 1244;
Best Local Similarity 100.0%; Pred. No. 6.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRHHCKSKAKRSRH 16

DB 827 RRHHCKSKAKRSRH 841

```

RESULT 2
Q8CH1 1
ID Q8CH1 PRELIMINARY; PRT; 1271 AA.
AC Q8CH1
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to period homolog 1 (Drosophila).
GN PER1
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver;
RA Strausberg R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC039768; AAH39768.1; -.
DR PIR; PT0697; PT0697.
DR MGD; MGI:1098283; Peri.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR000014; PAS_domain.
DR Pfam; PF00989; PAS; 1.
DR SMART; SM00091; PAS; 1.
DR PROSITE; PS01112; PAS; 1.
SQ SEQUENCE 1271 AA; 134158 MW; ACF1F27DFA6621CF CRC64;

Query Match 95.7%; Score 88; DB 11; Length 1271;
Best Local Similarity 100.0%; Pred. No. 6.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRHCRSKAKRSRH 16
DB 308 RRHCRSKAKRSRH 822

RESULT 3
Q8K3T3 1
ID Q8K3T3 PRELIMINARY; PRT; 1285 AA.
AC Q8K3T3
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Period 1 protein.
GN PER1.
OS Spalax judaei.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Spalacinae;
OC Spalax.
OX NCBI_TaxID=134510;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22199892; PubMed=12193657;
RA Avivi A., Oster H., Joel A., Albrecht U., Nevo E.;
RT "The molecular circadian clock in a blind mammal: three period
RT homologs in blind, subterranean mole rat."
RL Proc. Natl. Acad. Sci. U.S.A. 99:11718-11723 (2002).
DR EMBL; AJ345059; CAC95146.1; -.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR000014; PAS_domain.
DR Pfam; PF00989; PAS; 1.
DR SMART; SM00091; PAS; 2.
DR PROSITE; PS01112; PAS; 1.
SQ SEQUENCE 1285 AA; 135987 MW; D0494840FE9828D1 CRC64;

Query Match 95.7%; Score 88; DB 11; Length 1285;
Best Local Similarity 100.0%; Pred. No. 6.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 RRHCRSKAKRSRH 16
DB 824 RRHCRSKAKRSRH 838

RESULT 4
Q8SX50 1
ID Q8SX50 PRELIMINARY; PRT; 341 AA.
AC Q8SX50
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE RE04530p (CG9018-PB).
GN CG9018.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champagne M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Murtulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy L., Murphy D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskae R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhou L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RT Science 287:2185-2195 (2000).
RN [3]
RP SEQUENCE FROM N.A.

```

RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Ananides P.G., Brandon R.C., Rogers Y.,
 RA Banon J., An H., Baldwin D., Banon J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champagne M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Drenek D., Farfan D.,
 RA Ferreria S., Frise E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houch J., Hoskins R.A., Hostin D., Howard T.J.,
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
 RA Paclet J., Paragas V., Park S., Patel S., Pfeiffer B.,
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Swirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RT "Sequencing of *Drosophila melanogaster* genome.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Bernan B., Carlson J.W., Celniker S.E.,
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
 RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
 RT "Annotation of *Drosophila melanogaster* genome.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RL [6]
 RP SEQUENCE FROM N.A.
 RA FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY094845; AAM11198.1; -;
 DR EMBL; AE003474; AAN2216.1; -;
 DR FlyBase; FBgn0035318; CG9018.
 DR InterPro; IPR006903; DUF618.
 DR InterPro; IPR006569; RPR.
 DR Pfam; PF04818; DUF618; 1.
 DR SMART; SMO0582; RPR; 1.
 SQ SEQUENCE 341 AA; 38386 MW; 9B02FE4C95E75EC7 CRC64;
 Query Match 56.0%; Score 51.5; DB 5; Length 341;
 Best Local Similarity 68.8%; Pred. No. 1.6;
 Matches 11; Conservative 2; Mismatches 2; Indels 1; Gaps 1;
 QY 2 RRHCR-SKAKRSRH 16
 DB 173 RKEDRHRSKRSRH 188
 ID Q9W039 PRELIMINARY; PRT; 375 AA.
 AC Q9W039;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE CG9018 protein.
 GN CG9018.
 OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriidae; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananides P.G., Scher S.E., Li P.W., Hopkins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merklov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003474; AAF47620.1; -;
 DR FlyBase; FBgn0035318; CG9018.
 DR InterPro; IPR006903; DUF618.
 DR InterPro; IPR006569; RPR.
 DR Pfam; PF04818; DUF618; 1.
 DR SMART; SMO0582; RPR; 1.
 SQ SEQUENCE 375 AA; 41974 MW; 311062E3FB237AC5 CRC64;
 Query Match 56.0%; Score 51.5; DB 5; Length 375;
 Best Local Similarity 68.8%; Pred. No. 1.8;
 Matches 11; Conservative 2; Mismatches 2; Indels 1; Gaps 1;
 QY 2 RRHCR-SKAKRSRH 16
 DB 173 RKEDRHRSKRSRH 188
 ID Q8G4B1 PRELIMINARY; PRT; 179 AA.
 AC Q8G4B1;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 GN BL1478.
 OS Bifidobacterium longum.
 OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
 OC Bifidobacteriaceae; Bifidobacterium.
 OX NCBI_TaxID=216916;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCC 2705;
 RX MEDLINE=22294977; PubMed=12381787;
 RA Schell M.A., Karmirantzou M., Shel B., Vilanova D., Berger B.,
 RA Fessi G., Zwahlen M.-C., Desiere F., Bork P., Delley M.,
 RA Pridmore R.D., Arigoni F.;
 RT "The genome sequence of *Bifidobacterium longum* reflects its adaptation

```

RT to the human gastrointestinal tract.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427 (2002).
DR EMBL; A014778; AAN25273.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 179 AA; 20185 MW; B9C5F758A8C763C CRC64;

Query Match 54.3%; Score 50; DB 16; Length 179;
Best Local Similarity 62.5%; Pred. No. 1.6;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 SRRHCRSKAKSRHH 16
Db 47 SRRKPCRAKSRHH 62

RESULT 7
ID Q9D3A7 PRELIMINARY; PRT; 162 AA.
AC Q9D3A7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE 6330409D20Rik protein.
GN 6330409D20Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
RC MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Mateu Y., Nikaudo I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzaletti J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
DR EMBL; AK018150; BAB31096.1; -.
DR MGD; MGI:1917980; 6330409D20Rik.
SQ SEQUENCE 162 AA; 17986 MW; 925EB593D661FAC1 CRC64;

Query Match 53.3%; Score 49; DB 11; Length 162;
Best Local Similarity 53.8%; Pred. No. 2.1;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 4 HCRSKAKSRHH 16
Db 74 HCRRRRRHHH 86

RESULT 8
ID Q40367 PRELIMINARY; PRT; 325 AA.
AC Q40367;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Peroxidase precursor (Fragment).

```

```

GN PXDD.
OS Medicago sativa (Alfalfa).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
OX NCBI_TaxID=3879;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Siriver;
RA Abrahams S.L., Hayes C.M., Watson J.M.;
RT "Organ-specific expression of three peroxidase-encoding cDNAs from
RT lucerne (Medicago sativa).";
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; L36158; AAB41812.1; -.
DR PIR; T09667; T09667.
DR HSPP; P00433; ZATU.
DR GO; GO:0004601; P:peroxidase activity; IEA.
DR GO; GO:0006979; P:response to oxidative stress; IEA.
DR InterPro; IPR002016; Peroxidase.
DR Pfam; PF00141; peroxidase; 1.
DR PRINTS; PR00458; PEROXIDASE.
DR PROSITE; PS00435; PEROXIDASE_1; 1.
DR PROSITE; PS00436; PEROXIDASE_2; 1.
DR PROSITE; PS50873; PEROXIDASE_4; 1.
KW Signal.
FT SIGNAL 1 1 22 POTENTIAL.
FT CHAIN 23 325 PEROXIDASE.
FT CHAIN 23 325
SQ SEQUENCE 325 AA; 35931 MW; F7B9681D6AB3A9E CRC64;

Query Match 53.3%; Score 49; DB 10; Length 325;
Best Local Similarity 61.5%; Pred. No. 4;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 4 HCRSKAKSRHH 16
Db 16 HCRTRHQLSRHH 28

RESULT 9
ID Q95MF0 PRELIMINARY; PRT; 499 AA.
AC Q95MF0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Estrogen receptor beta 2.
GN ERBETA2.
OS Macaca arctoides (Stump-tailed macaque).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9540;
RN [1]
RP SEQUENCE FROM N.A.
RA Scobie G.A., Wilson J.A., Millar M.R., Macpherson S., Saunders P.T.;
RT "The estrogen receptor beta variant ERbeta2 is expressed in
RT a wide range of tissues in both Old and New World primates.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
DR EMBL; AF393815; AAK71317.1; -.
DR GO; GO:0005634; C:nucleus; ISS.
DR GO; GO:0030284; F:estrogen receptor activity; ISS.
DR GO; GO:0048019; F:receptor antagonist activity; ISS.
DR GO; GO:0005496; F:steroid binding; ISS.
DR GO; GO:0030520; P:estrogen receptor signaling pathway; ISS.
DR GO; GO:0030308; P:negative regulation of cell growth; ISS.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; ISS.
DR InterPro; IPR000536; Hormone_rec_lig.
DR InterPro; IPR001723; Sterhmrn_receptor.
DR InterPro; IPR008946; Str_ncl_receptor.
DR InterPro; IPR001628; Znf_C4steroid.

```

Pfam; PF00104; hormone_rec; 1.
 DR Pfam; PF00105; zf-C4; 1.
 DR PRINTS; PR00398; STRDHORMONER.
 DR PRINTS; PR00047; STROIDFINGER.
 DR ProDom; PD000035; Znf_C4steroid; 1.
 DR SMART; SM00430; HOLI; 1.
 DR SMART; SM00399; Znf_C4; 1.
 DR PROSITE; PS00031; NUCLEAR RECEPTOR; 1.
 DR DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
 KW Transcription regulation; Zinc; Zinc-finger.
 SQ SEQUENCE 499 AA; 55782 MW; A91DA345C818COA CRC64;
 Query Match 52.2%; Score 48; DB 6; Length 499;
 Best Local Similarity 75.0%; Pred. No. 8.7;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 5 HCRSKAKRSRHH 16
 DB 240 HCASKAKRSGSH 251
 RESULT 10
 Q8LM94 PRELIMINARY; PRT; 129 AA.
 ID Q8LM94
 AC Q8LM94
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN OSJNB0079801.8
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA McCombie W.R., de la Bastide M., Spiegel L., Preston R., Ferraro K.,
 RA Kuit K., Nascimento L., Zutavern T., Ballja V., Bell M., Baker J.,
 RA Mille B., Katzenberger F., Muller S., King L., Sullivan P., Yang C.,
 RA Dike S., O'Shaughnessy A., Palmer L., Dedhia N.;
 RT "Genomic sequence for Oryza sativa, Nipponbare strain, clone
 RT OSJNB0079801, from chromosome 10, complete sequence."
 RN Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA The Rice Chromosome 10 Sequencing Consortium;
 RT "In-depth view of structure, activity, and evolution of rice
 RT chromosome 10."
 RL Science 300:1566-1569(2003).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
 RA Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AC119149; BAM74392.1;
 DR EMBL; AB017074; AAP52888.1;
 DR Gramene; Q8LM94;
 KW Hypothetical protein.
 SQ SEQUENCE 129 AA; 13419 MW; 5ED75BC3186EAB CRC64;
 Query Match 51.1%; Score 47; DB 10; Length 129;
 Best Local Similarity 69.2%; Pred. No. 3.5;
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 4 HCRSKAKRSRHH 16
 DB 86 HAVRVNKAHSRHH 98
 RESULT 11

Q95Q04 PRELIMINARY; PRT; 340 AA.
 ID Q95Q04
 AC Q95Q04
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Y66D12A.11 protein.
 GN Y66D12A.11.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Feloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sulston J.E.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL; AL161712; CAC70133.1;
 DR WormPep; Y66D12A.11; CE28793.
 SQ SEQUENCE 340 AA; 38174 MW; 0DEE24A74BC7DAF3 CRC64;
 Query Match 51.1%; Score 47; DB 5; Length 340;
 Best Local Similarity 53.3%; Pred. No. 8.8;
 Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 QY 2 RRHCKRSKAKRSRHH 16
 DB 123 RRHKSQSNKRRHH 137
 RESULT 12
 O30258 PRELIMINARY; PRT; 375 AA.
 ID O30258
 AC O30258
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Coenzyme PQQ synthesis protein (PQQE).
 GN AF2413
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 OC Archaeoglobaceae; Archaeoglobus.
 OX NCBI_TaxID=2234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Richard K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Fleischmann R.D., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
 RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon Archaeoglobus fulgidus."
 RL Nature 390:364-370(1997).
 DR EMBL; AE001109; AAB91253.1;
 DR PIR; F69551; F69551.
 DR TIGR; AF2413;
 DR InterPro; IPR006638; Elp3.
 DR InterPro; IPR007197; Radical SAM.
 DR Pfam; PF04055; Radical_SAM; I.

DR SMART; SM00729; ELP3; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 375 AA; 42026 MW; 53F0E3D45A0A9CDE CRC64;
 Query Match 51.1%; Score 47; DB 17; Length 375;
 Best Local Similarity 72.7%; Pred. No. 9.7;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 5 HCRSKAKRSRH 15
 |||||
 DB 21 HCRKAIRKXH 31
 RESULT 13
 Q9LN46 ID Q9LN45 PRELIMINARY; PRT; 840 AA.
 AC Q9LN45; ID Q9LN45; PRT; 840 AA.
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE F18014.24
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
 RA Kim C., Altafi H., Bei Q., Chin C., Chlou J., Choi E., Conn L.,
 RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
 RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
 RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
 RA Ecker J.R.;
 RA "Genomic sequence for Arabidopsis thaliana BAC F18014 from chromosome
 RT I.";
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AC025808; AAF79430.1; -
 DR InterPro; IPR001810; F-box.
 DR Pfam; PF006652; Kelch_rep.
 DR Pfam; PF00646; F-box; 2.
 DR Pfam; PF01344; Kelch; 4.
 DR SMART; SM00612; Kelch; 3.
 SQ SEQUENCE 840 AA; 95576 MW; 293C6E64AB379CD6 CRC64;
 Query Match 51.1%; Score 47; DB 10; Length 840;
 Best Local Similarity 58.8%; Pred. No. 21;
 Matches 10; Conservative 1; Mismatches 4; Indels 2; Gaps 1;
 QY 2 RRHCRS--KAKRSRH 16
 |||||
 DB 178 RTHKCRSLPKWGRGRYH 194
 RESULT 14
 Q7SXU9 ID Q7SXU9 PRELIMINARY; PRT; 347 AA.
 AC Q7SXU9; ID Q7SXU9; PRT; 347 AA.
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Splicing factor.
 OS Brachydanio rerio (zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AB; TISSUE=Body;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raza S.S., Loquellano N.J., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmitz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AB; TISSUE=Body;
 RA Strausberg R.;
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC055238; AAH5238.1; -
 SQ SEQUENCE 347 AA; 40294 MW; C57ECEA7BDB55CC7 CRC64;
 Query Match 50.5%; Score 46.5; DB 13; Length 347;
 Best Local Similarity 66.7%; Pred. No. 11;
 Matches 10; Conservative 3; Mismatches 1; Indels 1; Gaps 1;
 QY 1 SRRHCRSKA-KRSR 14
 |||||
 DB 196 SRRHCRSGRRSR 210
 RESULT 15
 Q7UUR6 ID Q7UUR6 PRELIMINARY; PRT; 135 AA.
 AC Q7UUR6; ID Q7UUR6; PRT; 135 AA.
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN RB3125.
 OS Rhodospirillum rubrum.
 OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
 OC Planctomycetaceae; Pirellula.
 OX NCBI_TaxID=117;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1;
 RX MEDLINE=22735913; PubMed=12835416;
 RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
 RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
 RA Schleener H., Amann R., Reinhardt R.;
 RT "Complete genome sequence of the marine planctomycete Pirellula sp.
 RT strain 1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
 RL EMBL; BX294138; CAD73011.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 135 AA; 14862 MW; 83B9F9FCBA53DFF2 CRC64;
 Query Match 50.0%; Score 46; DB 16; Length 135;
 Best Local Similarity 61.5%; Pred. No. 5.3;
 Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 4 HCRSKAKRSRH 16
 |||||
 DB 50 HCRNATRSFHH 62
 Search completed: May 21, 2004, 12:54:45

Job time : 42 secs
